

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:51:32 : Search time 85 Seconds  
(without alignments)  
760.020 Million cell updates/sec

Title: US-10-017-372E-37

Perfect score: 2159

Sequence: 1 MAPSGLRLPLLLPLLLLV.....GRPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2048.5	94.9	390	AAE13596	Porcine transform
2	1926.5	89.2	390	AAAM39186	Human polypeptide
3	1919.5	88.9	390	7 AAF61468	PreTGF-beta gene p
4	1919.5	88.9	390	11 AAR04034	Sequence of pre-TG
5	1919.5	88.9	390	11 AAR05258	Human pre-transfor
6	1919.5	88.9	390	12 AAR13813	Human pro-TGF-beta
7	1919.5	88.9	390	16 AAR73596	Human TGF-beta 1 p
8	1919.5	88.9	390	17 AAR90827	Human transforming g
9	1916.5	88.8	390	23 AAU77101	Human transforming

10	1916.5	88.8	390	23	AAE16943	Human transforming
11	1916	88.7	391	24	ABB82780	TGFβ1 Arg25Pro pol
12	1914.5	88.7	390	13	AAE20124	Sequence of simian
13	1910.5	88.5	390	15	AAE46227	Human pre-TGF-beta
14	1909	88.4	391	16	AAE83034	Transforming growt
15	1908.5	88.4	390	19	AAW78785	Human pre-transfor
16	1907.5	88.4	390	22	AAE84601	Nucleotide sequenc
17	1905	88.2	391	9	AAE81362	Human transforming
18	1904	88.2	434	11	AAE03743	Monkey transforming
19	1900.5	88.0	390	24	ABB82781	TGFβ1 Arg25Pro pol
20	1883.5	87.2	385	11	AAE05663	Simian Transformin
21	1878	87.0	387	11	AAE05664	Simian Transformin
22	1867.5	86.5	390	11	AAE05452	Chimeric simian TG
23	1863.5	86.3	390	13	AAE27522	TGF-beta 1/beta 2
24	1844	85.4	389	13	AAE29657	TGF-beta 1. Homo
25	1839	85.2	453	22	AAE40972	Human polypeptide
26	1758.5	81.4	390	13	AAE20126	Sequence of hybrid
27	1757.5	81.4	390	11	AAE05749	Human TGF-beta2 ex
28	1751.5	81.1	390	11	AAE05665	Human transforming
29	1749.5	81.0	390	11	AAE05666	Hybrid transformin
30	1742	80.7	391	10	AAE91900	Sequence encoded b
31	1715	79.4	389	16	AAE79921	Simian-human hybr
32	1300	60.2	278	15	AAE53090	Polypeptide cross-
33	1295	60.0	278	12	AAE12541	Latency associated
34	1262.5	58.5	458	23	ABG1507	LAP-miFNB constru
35	1262.5	58.5	463	23	ABG31510	LAP-huIFNB constru
36	1169	54.1	290	22	ABG06792	Novel human diagno
37	1135	52.6	450	23	ABG31508	miFNB-LAP constru
38	1128	51.8	448	23	ABG31509	huIFNB-LAP constru
39	944	43.7	227	22	ABG20234	Novel human diagno
40	901	41.7	236	22	ABG20233	Novel human diagno
41	868.5	40.2	382	21	AAE08338	Amino acid sequenc
42	868.5	40.2	382	23	AAU77105	Frog transforming
43	832.5	38.6	456	13	AAE78786	Pig transforming g
44	828.5	38.4	412	16	AAE73598	Human TGF-beta 3 p
45	826	38.3	456	15	AAE46228	Pig TGF-beta-3. S

#### ALIGNMENTS

RESULT 1  
AAE13596  
ID AAE13596 standard; Protein: 390 AA.  
XX  
AC AAE13596;  
XX  
DT 26-FEB-2002 (first entry)  
XX  
DE Porcine transforming growth factor beta 1 (TGF-beta1) mutant.  
XX  
KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;  
KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;  
KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; mutant;  
KW mutain.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 223 /note= "Wild type Cys substituted with Ser"  
FT Misc-difference 225 /note= "Wild type Cys substituted with Ser"  
FT  
XX WO200181404-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US12980.  
XX  
PR 20-APR-2000; 2000US-199014P.  
XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Strober W, Nakamura K, Kitani A, Fuss LJ;

XX WPI: 2002-026155/03.

DR N-PSDB; AAD22696.

XX Composition for treating autoimmune diseases e.g. inflammatory bowel  
PT disease in humans, comprises vector containing transforming growth  
PT factor-beta under the control of inducible promoter

XX Example 1: Fig 1: 78pp: English.

XX The invention relates to a composition containing a vector comprising a  
CC gene encoding a regulatory transcription factor under the control of a  
CC promoter encoding a transforming growth factor-beta (TGF-beta). The  
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2  
CC or TGF-beta3, its variants or homologues, by transfecting a cell which  
CC is part of a host suspected of having an autoimmune disease, especially  
CC inflammatory bowel disease (IBD), under conditions such that the  
CC polypeptide encoded by the nucleic acid sequence in the vector is  
CC expressed. The vector is delivered using a delivery system. The delivery  
CC of the vector results in substantial elimination of symptoms of the  
CC autoimmune disease and increased production of IL-10 by the host. The  
CC composition is useful for treating various diseases with an autoimmune  
CC component such as multiple sclerosis, rheumatoid arthritis, systemic  
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis  
CC and psoriasis, and also for assaying the expression of a gene in a cell.  
CC The vector is further useful for screening of the effect of test  
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.  
CC The present sequence is porcine TGF-beta1 mutant.

XX Sequence 390 AA;

Query Match 94.9%; Score 2048.5; DB 23; Length 390;

Best Local Similarity 95.8%; Pred. No. 5.2e-171;

Matches 390; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLLLVLTGPPAAGLSTCKTIDMELVKRRKRIEIRGOILSKRLA 60  
DB 1 MAPSGRLRLPLLLPLLLVLTGPPAAGLSTCKTIDMELVKRRKRIEIRGOILSKRLA 60  
QY 61 SPSSQGDVPPGPPPLPAVALYNSTRDRVAGESVPEPEADYAKETRVLMVYESNQI 120  
DB 61 SPSSQGDVPPGPPPLPAVALYNSTRDRVAGESVPEPEADYAKETRVLMVYESNQI 120  
QY 121 YDKFKGTPLSLYMLFNTSELHEAVPEPVLSSRAELRLRLKVEQHVLYKQYNSDSWR 180  
DB 121 YDKFKGTPLSLYMLFNTSELHEAVPEPVLSSRAELRLRLKVEQHVLYKQYNSDSWR 180  
QY 181 YLSNRLAPSDSPWLSFDVTVGVVQWLTTRREAIEGFRLSAHSDDSKNTLHVEINGFN 240  
DB 181 YLSNRLAPSDSPWLSFDVTVGVVQWLTTRREAIEGFRLSAHSDDSKNTLHVEINGFN 240  
QY 241 SGRRGLATIHGMNRPFLLNATPLERAOHLHSSRRHALOTNSYPYDVPDYASLADTN 300  
DB 241 SGRRGLATIHGMNRPFLLNATPLERAOHLHSSRRHALOTNSYPYDVPDYASLADTN 300  
QY 301 YCFSSTEKNCCVQYLDYFRKDLGKWIHEPKGYHANFCLGCPYINSLDTQYSKVIALY 360  
DB 301 YCFSSTEKNCCVQYLDYFRKDLGKWIHEPKGYHANFCLGCPYINSLDTQYSKVIALY 360  
QY 361 NOHNPASAPCCVQYLDYFRKDLGKWIHEPKGYHANFCLGCPYINSLDTQYSKVIALY 360  
DB 361 NOHNPASAPCCVQYLDYFRKDLGKWIHEPKGYHANFCLGCPYINSLDTQYSKVIALY 360  
QY 344 NOHNPASAPCCVQYLDYFRKDLGKWIHEPKGYHANFCLGCPYINSLDTQYSKVIALY 390  
DB 344 NOHNPASAPCCVQYLDYFRKDLGKWIHEPKGYHANFCLGCPYINSLDTQYSKVIALY 390

RESULT 2

AA39186

ID AAM39186 standard; Protein; 390 AA.

XX AAM39186;

XX AAM39186;

XX AAM39186;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2331.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 23-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI58342.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries

XX Example 4; SEQ ID NO 2331; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nontropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localized neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX Sequence 390 AA;

XX Query Match 89.2%; Score 1926.5; DB 22; Length 390;

XX Best Local Similarity 89.9%; Pred. No. 2.6e-160;

XX Matches 366; Conservative 10; Mismatches 14; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLLLVLTGPPAAGLSTCKTIDMELVKRRKRIEIRGOILSKRLA 60

DB 1 MPSPGRLRLPLLLPLLLVLTGPPAAGLSTCKTIDMELVKRRKRIEIRGOILSKRLA 60

QY 61 SPSSQGDVPPGPPPLPAVALYNSTRDRVAGESVPEPEADYAKETRVLMVYESNQI 120

DB 61 SPSSQGDVPPGPPPLPAVALYNSTRDRVAGESVPEPEADYAKETRVLMVYESNQI 120

QY 121 YDKFKGTSPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKQHVLYQKYSNDSWR 180  
DB 121 YDKFKGTSPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKQHVLYQKYSNDSWR 180  
QY 181 YLSNRLAPSDSPWLSFCTGTGVVROWLTREALEGFRLSAHSDDSKDNTLHVEINGFN 240  
DB 181 YLSNRLAPSDSPWLSFCTGTGVVROWLTREALEGFRLSAHSDDSKDNTLHVEINGFN 240  
QY 241 SGRGDLATIHGMNRPFLILMATPLERAQHLHSSRRRALDNTSYDVPDYASLALDTN 300  
DB 241 TGRGDLATIHGMNRPFLILMATPLERAQHLHSSRRRALDNTSYDVPDYASLALDTN 300  
QY 301 YCFSTERNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVIALY 360  
DB 284 YCFSTERNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVIALY 343  
QY 361 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
DB 344 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 3  
AAP61468  
ID AAP61468 standard; Protein; 390 AA.  
AC AAP61468;  
XX  
DT 31-OCT-2002 (updated)  
DT 28-OCT-1991 (first entry)  
DE PrefGF-beta gene product.  
KW Transforming growth factor beta; cancer; wound healing.  
XX Unidentified.  
OS  
FH Key Location/Qualifiers  
FT Protein 279..390  
XX  
PN EP200341-A.  
XX  
PD 10-DEC-1986.  
XX  
PF 21-MAR-1986; 86EP-0302112.  
XX  
PR 22-MAR-1985; 85US-0715142.  
PR 13-MAR-1987; 87US-0025423.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Derynck RMA;  
XX  
DR WPI; 1986-326875/50.  
DR N-PSDB; AAN60572.  
XX  
PT TGF-beta prodn. from transformed hosts - useful esp. for treating  
PT wounds (J6 2/9/86).  
XX  
PS Disclosure; Fig 1b; 26pp; English.  
XX  
CC The gene product is known to stimulate cell proliferation and  
CC inhibit anchorage-dependent growth of a variety of human cancer cell  
CC lines, it is esp. useful in treatment of burns and the promotion of  
CC surface and internal wound healing. TGF-beta may be expressed from a  
CC transformed CHO cell line.  
CC (Updated on 31-OCT-2002 to add missing OS field.)  
XX  
SQ Sequence 390 AA;

Query Match 88.9%; Score 1919.5; DB 7; Length 390;  
Best Local Similarity 89.7%; Pred. No. 1.1e-159;  
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLMLLVLTGCPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60  
DB 1 MPPSGRLRLPLLLPLMLLVLTGCPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60  
QY 61 SPPSGDVPVPPGLPEAVLALYNSTRDRVAGSEPEPEADYYAKVETRVLMVYESGNOI 120  
DB 61 SPPSGDVPVPPGLPEAVLALYNSTRDRVAGSEPEPEADYYAKVETRVLMVETHNEI 120  
QY 121 YDKFKGTSPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKQHVLYQKYSNDSWR 180  
DB 121 YDKFKGTSPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKQHVLYQKYSNDSWR 180  
QY 181 YLSNRLAPSDSPWLSFCTGTGVVROWLTREALEGFRLSAHSDDSKDNTLHVEINGFN 240  
DB 181 YLSNRLAPSDSPWLSFCTGTGVVROWLTREALEGFRLSAHSDDSKDNTLHVEINGFN 240  
QY 241 SGRGDLATIHGMNRPFLILMATPLERAQHLHSSRRRALDNTSYDVPDYASLALDTN 300  
DB 241 TGRGDLATIHGMNRPFLILMATPLERAQHLHSSRRRALDNTSYDVPDYASLALDTN 300  
QY 301 YCFSTERNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVIALY 360  
DB 284 YCFSTERNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVIALY 343  
QY 361 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
DB 344 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4  
AAR04034  
ID AAR04034 standard; protein; 390 AA.  
XX  
AC AAR04034;  
XX  
DT 25-MAR-2003 (updated)  
DT 31-OCT-2002 (updated)  
DT 31-MAY-1989 (first entry)  
XX  
DE Sequence of pre-TGF-beta 1.  
XX  
KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth  
KW inhibition.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 348..500  
XX  
PN W08912101-A.  
XX  
PD 14-DEC-1989.  
XX  
PF 08-JUN-1988; 88WO-US01945.  
XX  
PR 08-JUN-1988; 88WO-US01945.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Derynck RMA, Gooddel DV;  
XX  
DR WPI; 1990-007474/01.  
DR N-PSDB; AAQ02815.  
XX  
PT Nucleotide sequence encoding transforming growth factor beta-3 used as a  
PT probe, or to produce TGF beta 3, for inhibiting growth of certain normal  
PT and neoplastic cells, eg A549.  
XX  
PS Disclosure; Fig. 2; 61pp; English.  
XX  
CC Sequence is an exon of transforming growth factor-beta 1 (pre-TGF-beta  
CC 1) polypeptide and corresponds to AA's 288-338 of mature TGF-beta 1. The  
CC nucleic acid encoding second subtype of TGF-beta (TGF-beta 3) is useful

CC as a probe or to produce TGF-beta 3 for inhibition of normal and  
 CC neoplastic cell growth.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX Sequence 390 AA:

```

Query Match      88.9%; Score 1919.5; DB 11; Length 390;
Best Local Similarity 89.7%; Pred. No. 1.1e-159;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKKRIEAIHQILSKRLA 60
DB 1 MPSPGRLRLPLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKKRIEAIHQILSKRLA 60
QY 61 SPSPQGDVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKEVTRVLMVTHNEI 120
DB 61 SPSPQGEVPPGPLEAVLALYNSTRDRVAGESAEPPEPEADYAKEVTRVLMVTHNEI 120
QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
DB 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
QY 181 YLSNRLAPSDSPSEWLSFDVTGVVQWLTRREAIEGFRLSAHSSDSKDNTHLVEINGFN 240
DB 181 YLSNRLAPSDSPSEWLSFDVTGVVQWLTRREAIEGFRLSAHSSDSKDNTHLVEINGFN 240
QY 241 SGRGDLATIHGMNRPFLLMATPLERAQHLHSSRRALDNTNSYPDVPDYASLALDTN 300
DB 241 TGRGDLATIHGMNRPFLLMATPLERAQHLHSSRRALDNTNSYPDVPDYASLALDTN 300
QY 301 YCFSTSEKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
DB 284 YCFSTSEKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
QY 361 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVQELSNMIVRSCKS 407
DB 344 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVQELSNMIVRSCKS 390

RESULT 5
AAR05258
ID AAR05258 standard; protein; 390 AA.
AC AAR05258;
XX
XX 25-MAR-2003 (updated)
DT 05-AUG-1990 (first entry)
XX
XX Human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).
DE Transforming growth factor-beta-1 (TGF-beta-1);
KW neoplastic cell line inhibition;
KW EGF-potentiated anchorage-independent growth.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..278
FT Protein 279..2011
FT Domain 8..23
FT Modified-site /note="hydrophobic domain" 82..84
FT Modified-site /note="potential N-glycosylation site" 136..138
FT Modified-site /note="as above" 176..178
FT Modified-site /note="as above" 277..278
FT Cleavage-site /note="proteolytic cleavage site"
XX
XX US4886747-A.

```

```

XX 12-DEC-1989.
XX
XX 13-MAR-1987; 87US-0025423.
XX
XX 13-MAR-1987; 87US-0025423.
XX
XX 22-MAR-1985; 85US-0715142.
XX
XX (GETH ) GENENTECH INC.
XX
XX Derynck RMA, Goeddel DV;
XX
XX WPI: 1990-051338/07.
XX N-PSDB; AAQ93301.
XX
XX Nucleic acid encoding transforming growth factor-beta -
XX cloned into expression vectors for expression in eukaryotic host
XX cells for therapeutic use
XX
XX Disclosure; Fig 1b; 28pp; English.
XX
XX Pre-TGF-beta-1 is a polypeptide of 390 amino acids. Post-translational
XX cleavage of the precursor gives rise to the mature TGF-beta monomer.
XX The sequence for human TGF-beta was determined by direct amino acid
XX sequence analysis and by deduction from the TGF-beta cDNA. It is
XX capable of inducing EGF-potentiated anchorage-independent growth of
XX target cell lines, and/or growth inhibition of neoplastic cell lines. It
XX can be used for treating wounds, eg burns or epidermal ulcers.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 390 AA:
Query Match      88.9%; Score 1919.5; DB 11; Length 390;
Best Local Similarity 89.7%; Pred. No. 1.1e-159;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKKRIEAIHQILSKRLA 60
DB 1 MPSPGRLRLPLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKKRIEAIHQILSKRLA 60
QY 61 SPSPQGDVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKEVTRVLMVTHNEI 120
DB 61 SPSPQGEVPPGPLEAVLALYNSTRDRVAGESAEPPEPEADYAKEVTRVLMVTHNEI 120
QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
DB 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
QY 181 YLSNRLAPSDSPSEWLSFDVTGVVQWLTRREAIEGFRLSAHSSDSKDNTHLVEINGFN 240
DB 181 YLSNRLAPSDSPSEWLSFDVTGVVQWLTRREAIEGFRLSAHSSDSKDNTHLVEINGFN 240
QY 241 SGRGDLATIHGMNRPFLLMATPLERAQHLHSSRRALDNTNSYPDVPDYASLALDTN 300
DB 241 TGRGDLATIHGMNRPFLLMATPLERAQHLHSSRRALDNTNSYPDVPDYASLALDTN 300
QY 301 YCFSTSEKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
DB 284 YCFSTSEKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
QY 361 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVQELSNMIVRSCKS 407
DB 344 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVQELSNMIVRSCKS 390

RESULT 6
AAR13813
ID AAR13813 standard; Protein; 390 AA.
XX AAR13813;
XX
XX 20-NOV-1991 (first entry)
XX
XX

```



```

DE XX Human pro-TGF-beta 1.
KW XX Osteogenetic; tumoricidal.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Peptide 1..29
FT Peptide /note= "signal peptide"
FT Peptide 30..390
FT Peptide /note= "pro-TGF-beta 1"
FT Peptide 279..390
FT Peptide /note= "TGF beta 1"
XX JP03180192-A.
XX PD 06-AUG-1991.
XX PF 07-DEC-1989; 89JP-0318243
XX PR 07-DEC-1989; 89JP-0318243.
XX PA (KIRI ) KIRIN BREWERY KK.
XX DR WPI: 1991-271579/37.
XX DR N-PSDB; AAQ13392.
XX PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by
PT preparing DNA chain contg. base sequence coding for human
PT pre-pro-TGF-beta 1, forming expression vector etc.
XX PS Claim 1; Fig 1; 16pp; Japanese.
XX CC The amino acid sequence codes for human prepro-TGF-beta 1 which
CC can be produced by recombinant methods, it has osteogenetic and
CC tumoricidal activity.
XX SQ Sequence 390 AA;

Query Match 88.9%; Score 1919.5; DB 12; Length 390;
Best Local Similarity 89.7%; Pred. No. 1.1e-159;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGLRLPLLLPLLMWLVLTGPPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
DB 1 MPSPGLRLPLLLPLLMWLVLTGPPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
QY 61 SPSPQGVPPGPPPLPEAVLYLNSTRDVAGESVEPEPEADYYAKEVTRVLMVSGNQI 120
DB 61 SPSPQGVPPGPPPLPEAVLYLNSTRDVAGESVEPEPEADYYAKEVTRVLMVTHNEI 120
QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
DB 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
QY 181 YLSNRLAPSDSPWLSFVDTGVVROWLTPRRAIEGFRLSAHSSSDSKDNTLHVINGFN 240
DB 181 YLSNRLAPSDSPWLSFVDTGVVROWLTPRRAIEGFRLSAHSSSDSKDNTLHVINGFN 240
QY 241 SGRRGDLATHGNRPFLIMATPLERAQHLHSSRRALDNTNSYPDVPDYLALDTN 300
DB 241 TGRRGDLATHGNRPFLIMATPLERAQHLHSSRRALDNTNSYPDVPDYLALDTN 300
QY 301 YCFSSTEKNCVQRYIDF RKDLGWKWIHEPKGYHANFCLGCPPIYIWSLDQYSKVLALY 360
DB 284 YCFSSTEKNCVQRYIDF RKDLGWKWIHEPKGYHANFCLGCPPIYIWSLDQYSKVLALY 343
QY 361 NQHPGASAPCCVQALPEPLPIYIVYGRKPKVEQLSNMIVRSCKS 407
DB 344 NQHPGASAPCCVQALPEPLPIYIVYGRKPKVEQLSNMIVRSCKS 390

RESULT 7

```

---

```

AAR73596
ID AAR73596 standard: Protein: 390 AA.
XX AC AAR73596;
XX DT 25-MAR-2003 (updated)
XX DT 20-DEC-1995 (first entry)
XX DE Human TGF-beta 1 protein.
XX KW Transforming growth factor-beta: Human TGF-beta protein; TGF-beta 1;
KW TGF-beta 3; osteogenic cell source; OCS; bone deficiency;
KW bone-inducing cofactor.
XX OS Homo sapiens.
XX PN US5409896-A.
XX PD 25-APR-1995.
XX PF 12-NOV-1993; 93US-0132405.
XX PR 01-SEP-1989; 89US-0401906.
XX PR 12-NOV-1991; 91US-0790856.
XX PR 18-MAY-1993; 93US-0063841.
XX PR 12-NOV-1993; 93US-0132405.
XX PA (GETH ) GENENTECH INC.
XX PI Ammann AJ, Rudman CG;
XX WI WPI; 1995-169610/22.
XX PT Compsn. for treating skeletal tissue deficiency - comprising
XX PT transforming growth factor-beta and an osteogenic cell source in a
XX PT carrier
XX PS Claim 3; Column 15-18; 19pp; English.
XX CC This sequence represents human transforming growth factor-beta 1
XX CC (TGF-beta 1). This sequence and the sequence for human TGF-beta 3 (see
XX CC AAR73598) are claimed within the scope of the invention. The invention
XX CC is a composition consisting of a TGF-beta protein and an osteogenic cell
XX CC source (OCS) formulated in an acceptable carrier other than a bone
XX CC morphogenic cofactor. This composition can be used for the restoration
XX CC of bone deficiency. This provides for the generation of mature bone
XX CC only where it is required, without the inclusion of a specific
XX CC bone-inducing cofactor. This method can be used with any of the 5 human
XX CC TGF-beta's or with TGF-beta from other species.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 390 AA;

Query Match 88.9%; Score 1919.5; DB 16; Length 390;
Best Local Similarity 89.7%; Pred. No. 1.1e-159;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGLRLPLLLPLLMWLVLTGPPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
DB 1 MPSPGLRLPLLLPLLMWLVLTGPPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
QY 61 SPSPQGVPPGPPPLPEAVLYLNSTRDVAGESVEPEPEADYYAKEVTRVLMVSGNQI 120
DB 61 SPSPQGVPPGPPPLPEAVLYLNSTRDVAGESVEPEPEADYYAKEVTRVLMVTHNEI 120
QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
DB 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180
QY 181 YLSNRLAPSDSPWLSFVDTGVVROWLTPRRAIEGFRLSAHSSSDSKDNTLHVINGFN 240
DB 181 YLSNRLAPSDSPWLSFVDTGVVROWLTPRRAIEGFRLSAHSSSDSKDNTLHVINGFN 240

```

QY 241 SGRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDVPDYASALDNTN 300  
 :|||||  
 Db 241 TCRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDVPDYASALDNTN 283  
 :|||||  
 QY 301 YCFSTKNCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTOYSKVLALY 360  
 :|||||  
 Db 284 YCFSTKNCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTOYSKVLALY 343  
 :|||||  
 QY 361 NQHPGASAPCCVPOALEPIPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 :|||||  
 Db 344 NQHPGASAPCCVPOALEPIPIVYVGRKPKVEQLSNMIVRSCKCS 390  
 :|||||

## RESULT 8

AAK90827  
 ID AAK90827 standard; Protein; 390 AA.

XX AAK90827;

XX 25-MAR-2003 (updated)

DT 25-JAN-1980 (first entry)

XX Pre-transforming growth factor beta 1.

XX transforming growth factor beta 1; wound healing;  
 KW recombinant production.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 8..23

FT Modified-site /note= "hydrophobic domain"

FT Modified-site /note= "potential N-glycosylation site"

FT Modified-site /note= "potential N-glycosylation site"

FT Modified-site /note= "potential N-glycosylation site"

FT Cleavage-site /note= "trypsin-like peptidase cleavage site"

FT Protein 279..390

FT /label= mature\_TGF\_beta\_1

FT US5482851-A.

XX 09-JAN-1996.

XX 05-NOV-1993; 93US-0147364.

XX 13-MAR-1987; 87US-0025423.

XX 22-MAR-1985; 85US-0715142.

XX 04-AUG-1989; 89US-0389929.

XX 04-MAR-1992; 92US-0845893.

XX 05-NOV-1993; 93US-0147364.

XX (GETH ) GENENTECH INC.

XX Derynck RMA, Goeddel DV;

XX WPI; 1996-076891/08.

XX N-PSDB; AAT15720.

XX New recombinant human transforming growth factor-beta prods. - produced  
 PT using Chinese hamster ovary cells, for use in diagnostic applications  
 PT or in therapy

XX Example 3; Fig 1A-C; 26pp; English.

XX The pre-transforming growth factor (TGF) beta 1 protein is encoded  
 CC by AAT15720. The mature TGF beta 1 monomer is cleaved from the  
 CC precursor at the Arg-Arg dipeptide immediately preceding the mature  
 CC TGF-beta 1 NH2-terminus. It does not contain a recognisable N-terminal  
 CC signal peptide typical of most secreted proteins. The pre-TGF beta 1

CC contains several pairs of basic residues which could undergo  
 CC post-translational cleavage and give rise to separate polypeptide  
 CC entities. The precursor contains 3 potential N-glycosylation sites, none  
 CC of which are localised in the mature TGF beta 1. This is useful in  
 CC purification of the mature protein. TGF beta 1 can be used in, e.g. wound  
 CC healing.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 390 AA;

Query Match 88.9%; Score 1919.5; Da 17; Length 390;

Best Local Similarity 89.7%; Pred. No. 1; 159;

Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGLRLPLLLPLLLVLTGPPAAGLSTCKTIDMELVKKRIEATRGQILSKRLA 60

Db 1 MPPSGLRLPLLLPLLLVLTGPPAAGLSTCKTIDMELVKKRIEATRGQILSKRLA 60

QY 61 SPDSQGVPPGLPEAVLALYNSTRVAGESVEPEPEADYYAKEVTRVLWVESGNQI 120

Db 61 SPDSQGVPPGLPEAVLALYNSTRVAGESVEPEPEADYYAKEVTRVLWVETHNEI 120

QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVQHVVELYQKYSNDSWR 180

Db 121 YDKFKGTTHSIYMFNTSELREAVPEPVLLSRALRLRLKLVQHVVELYQKYSNDSWR 180

QY 181 YLSNRLIAPSDPEWLSFDVTGVVROWLITRRAIEGFRLSAHSDDSKDNTLHVINGFN 240

Db 181 YLSNRLIAPSDPEWLSFDVTGVVROWLITRRAIEGFRLSAHSDDSKDNTLHVINGFT 240

QY 241 SGRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDVPDYASALDNTN 300

Db 241 TCRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDVPDYASALDNTN 283

QY 301 YCFSTKNCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTOYSKVLALY 360

Db 284 YCFSTKNCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTOYSKVLALY 343

QY 361 NQHPGASAPCCVPOALEPIPIVYVGRKPKVEQLSNMIVRSCKCS 407

Db 344 NQHPGASAPCCVPOALEPIPIVYVGRKPKVEQLSNMIVRSCKCS 390

## RESULT 9

AAU77101

ID AAU77101 standard; Protein; 390 AA.

XX AAU77101;

XX 05-JUN-2002 (first entry)

XX Human transforming growth factor beta 1 (TGF-beta-1) polypeptide.

XX Human; transforming growth factor beta; TGF-beta; insulin production;  
 KW type I diabetes mellitus; pancreatic cell outgrowth; wound healing;  
 KW pancreatic duct tissue; ischaemia; stroke; nervous system aging;  
 KW neurological condition; neurodegenerative disease; inflammation;  
 KW nasal injury; chemical injury; traumatic injury; tumour-induced injury;  
 KW amyotrophic lateral sclerosis; spinocerebellar degeneration;  
 KW immunological disease; multiple sclerosis; TGF-beta-1.

XX Homo sapiens.

XX WO200212336-A2.

XX 14-FEB-2002.

XX 09-FEB-2001; 2001WO-US04192.

XX 09-AUG-2000; 2000US-0635368.

XX (CURI-) CURIS INC.

PI Wang M, Pang K;  
 DR WPI; 2002-257468/30.  
 XX  
 PT Treating a subject with a disorder resulting from insufficient insulin  
 PT production, and inducing outgrowth of pancreatic cells, involves using  
 PT a transforming growth factor beta therapeutic -  
 XX  
 PS Disclosure; Fig 1; 77pp; English.  
 XX  
 CC The invention relates to treating a subject with a disorder resulting  
 CC from insufficient insulin production, involving contacting the subject  
 CC with a transforming growth factor beta (TGF-beta) therapeutic. TGF-beta  
 CC polypeptides can be used for treating a subject with a disorder resulting  
 CC from insufficient insulin production, e.g. type I diabetes mellitus, and  
 CC for inducing outgrowth of pancreatic cells associated with pancreatic  
 CC duct tissue within a subject. A composition comprising a TGF-beta protein  
 CC may be useful in wound healing and treatment of neurological conditions  
 CC derived from acute, subacute or chronic injury to the nervous system,  
 CC (including traumatic injury, chemical injury, vascular injury and deficits  
 CC such as ischemia resulting from stroke), together with  
 CC infectious/inflammatory and tumour-induced injury, aging of the nervous  
 CC system including Alzheimer's disease, chronic neurodegenerative diseases  
 CC including Parkinson's disease, Huntington's chorea, amyotrophic lateral  
 CC sclerosis, spinocerebellar degenerations and chronic immunological  
 CC diseases of the nervous system or affecting the nervous system, including  
 CC multiple sclerosis. This sequence represents the human TGF-beta-1  
 CC protein.  
 XX  
 SQ Sequence 390 AA:  
 Query Match 88.8%; Score 1916.5; DB 23; Length 390;  
 Best Local Similarity 89.7%; Pred. No. 1.9e-159;  
 Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;  
 QY 1 MAPSGRLRLPLLLPLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIRAEIRGOILSKRLA 60  
 Db 1 MPSPGLRLLLPLLLPLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIRAEIRGOILSKRLA 60  
 QY 61 SPSSQGVPPGPLEPNAVLVLYNSTDRVAGESVEPEPEADYAKETRVLMVSGNQI 120  
 Db 61 SPSSQGVPPGPLEPNAVLVLYNSTDRVAGESVEPEPEADYAKETRVLMVETHNEI 120  
 QY 121 YDKFKTPTSHLYMIFNTSELREAVPEPVLVLSRAELRLRLKLVQHVLYKQYNSDSWR 180  
 Db 121 YDKFKOSTHSIYMFNTSELREAVPEPVLVLSRAELRLRLKLVQHVLYKQYNSNSWR 180  
 QY 181 YLSNRLAPSDSPWLSFDVVTGWVQWLTREARICFRLSAHSDDSKDNTLHVINGFN 240  
 Db 181 YLSNRLAPSDSPWLSFDVVTGWVQWLTREARICFRLSAHSDDSKDNTLQVDINGFT 240  
 QY 241 SGRRGDLATHGNRPFLLLMATPLERAOHLHSSRRHRRALDNTNSYPDVPDYASIALDNTN 300  
 Db 241 TGRRGDLATHGNRPFLLLMATPLERAOHLHSSRRHRRALDNTNSYPDVPDYASIALDNTN 300  
 QY 301 YCFSTTEKNCVQRLYIDFR<DLGWKWIHEPKGYHANFCIGPCPIWMSLDTOYSKVLALY 360  
 Db 284 YCFSTTEKNCVQRLYIDFR<DLGWKWIHEPKGYHANFCIGPCPIWMSLDTOYSKVLALY 343  
 QY 361 NOHNPASAPCCVPOALEP<PIVYVYGRKPKVEQLSNMIVRSCKCS 407  
 Db 344 NOHNPASAPCCVPOALEP<PIVYVYGRKPKVEQLSNMIVRSCKCS 390  
 RESULT 10  
 ID AAE16943  
 XX AAE16943 standard; Protein; 390 AA.  
 AC AAE16943;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX Human transforming growth factor-beta (TGF-beta1) protein.

XX KW Human; transforming growth factor-beta1; TGF-beta1; osteoporosis;  
 KW latency associated peptide; LAP; integrin alphavbeta3; apoptosis;  
 KW immunomodulation; inflammatory disease; fibrotic disease; cancer;  
 KW diabetic retinopathy; chronic obstructive pulmonary disorder;  
 KW bone resorption; rheumatoid arthritis; psoriasis; restenosis;  
 KW atherosclerosis; liver fibrosis; asthma; cystostatic; osteopathic;  
 KW ophthalmological; arteriosclerotic; vasotropic.  
 XX OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..229  
 FT Region 30..278  
 FT Domain 244..246  
 FT Protein 279..390  
 FT /note= "Signal peptide"  
 FT /note= "LAP-beta1"  
 FT /note= "RGD motif"  
 FT /note= "Human mature TGF-beta1 protein"  
 XX WO200190186-A1.  
 XX 29-NOV-2001.  
 XX 25-MAY-2001; 2001WO-GB02352.  
 XX 26-MAY-2000; 2000GB-0012991.  
 XX 05-JAN-2001; 2001GB-0000286.  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Ludbrook S, Barry S, Horgan C, Miller D;  
 XX WPI; 2002-097645/13.  
 XX  
 PT Identifying modulators of interactions between latency associated  
 PT peptides and integrin alphavbeta3 for therapeutic use, by contacting the  
 PT peptide and integrin with a test product and determining if the product  
 PT modulates interaction -  
 XX  
 PS Disclosure; Page 37-39; 44pp; English.  
 XX  
 CC The invention relates to a method for identification of a modulator  
 CC of the interaction between latency associated peptide (LAP) of  
 CC transforming growth factor-beta1 (TGF-beta1) and integrin alphavbeta3.  
 CC The method is useful for identifying a modulator of the interaction  
 CC between LAP and integrin alphavbeta3. The method is useful for  
 CC immunomodulation, in the treatment of inflammatory disease, fibrotic  
 CC disease, cancer, diabetic retinopathy, bone resorption or osteoporosis,  
 CC and for preventing apoptosis administering the modulator to the host.  
 CC The modulator (inhibitor of the interaction between LAP-beta1 and  
 CC integrin alphavbeta3) is useful in the manufacture of a medicament for  
 CC immunomodulation. The modulator (activator of the interaction between  
 CC LAP-beta1 and integrin alphavbeta3) is useful in the manufacture of  
 CC medicament for preventing apoptosis. The modulator is useful for  
 CC treating a inflammatory or fibrotic disease such as chronic obstructive  
 CC pulmonary disorder, rheumatoid arthritis, psoriasis, restenosis,  
 CC atherosclerosis, liver fibrosis and asthma. The present sequence is  
 CC human TGF-beta1 protein.  
 XX  
 SQ Sequence 390 AA;  
 Query Match 88.8%; Score 1916.5; DB 23; Length 390;  
 Best Local Similarity 89.7%; Pred. No. 1.9e-159;  
 Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;  
 QY 1 MAPSGRLRLPLLLPLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIRAEIRGOILSKRLA 60  
 Db 1 MPSPGLRLLLPLLLPLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIRAEIRGOILSKRLA 60  
 QY 61 SPSSQGVPPGPLEPNAVLVLYNSTDRVAGESVEPEPEADYAKETRVLMVSGNQI 120  
 Db 61 SPSSQGVPPGPLEPNAVLVLYNSTDRVAGESVEPEPEADYAKETRVLMVETHNEI 120  
 QY 121 YDKFKTPTSHLYMIFNTSELREAVPEPVLVLSRAELRLRLKLVQHVLYKQYNSDSWR 180  
 Db 121 YDKFKOSTHSIYMFNTSELREAVPEPVLVLSRAELRLRLKLVQHVLYKQYNSNSWR 180  
 QY 181 YLSNRLAPSDSPWLSFDVVTGWVQWLTREARICFRLSAHSDDSKDNTLHVINGFN 240  
 Db 181 YLSNRLAPSDSPWLSFDVVTGWVQWLTREARICFRLSAHSDDSKDNTLQVDINGFT 240  
 QY 241 SGRRGDLATHGNRPFLLLMATPLERAOHLHSSRRHRRALDNTNSYPDVPDYASIALDNTN 300  
 Db 241 TGRRGDLATHGNRPFLLLMATPLERAOHLHSSRRHRRALDNTNSYPDVPDYASIALDNTN 300  
 QY 301 YCFSTTEKNCVQRLYIDFR<DLGWKWIHEPKGYHANFCIGPCPIWMSLDTOYSKVLALY 360  
 Db 284 YCFSTTEKNCVQRLYIDFR<DLGWKWIHEPKGYHANFCIGPCPIWMSLDTOYSKVLALY 343  
 QY 361 NOHNPASAPCCVPOALEP<PIVYVYGRKPKVEQLSNMIVRSCKCS 407  
 Db 344 NOHNPASAPCCVPOALEP<PIVYVYGRKPKVEQLSNMIVRSCKCS 390  
 RESULT 10  
 ID AAE16943  
 XX AAE16943 standard; Protein; 390 AA.  
 AC AAE16943;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX Human transforming growth factor-beta (TGF-beta1) protein.

Db 61 SPFGQVPPGLPEAVLALYNSTRDRVAGESNPEPEADYAKVETRVLMVETHNEI 120  
 QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRRAELRLRLKLVQEHVELYQKYSNDSWR 180  
 Db 121 YDKFKGTSTHSIYMFNTSELREAVPEPVLLSRRAELRLRLKLVQEHVELYQKYSNDSWR 180  
 QY 181 YLSNRLAPSDSPWLSFDVTGVVROWLTPREAEIEGFRLSAHSSSDKNTLHVEINGFN 240  
 Db 181 YLSNRLAPSDSPWLSFDVTGVVROWLTPREAEIEGFRLSAHSSSDKNTLHVEINGFN 240  
 QY 241 SGRGDLATIHGMNRPFLLMATPLERAOHLHSSRRRALDNTNSYDVPDVASLALDTN 300  
 Db 241 TGRGDLATIHGMNRPFLLMATPLERAOHLHSSRRRALDNTNSYDVPDVASLALDTN 283  
 QY 301 YCFSSSTEKNCCVQOLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVIALY 360  
 Db 284 YCFSSSTEKNCCVQOLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVIALY 343  
 QY 361 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 Db 344 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390  
 RESULT 11  
 ABB82780  
 ID ABB82780 standard; Protein: 391 AA.  
 XX  
 AC ABB82780;  
 XX  
 DT 18-MAR-2003 (first entry)  
 XX  
 DE TGFβ1 Arg25Pro polymorphism G-allele protein sequence.  
 XX  
 KW Polymorphism; transforming growth factor beta 1; TGFβ1; TGFβ-1;  
 KW renal failure; nephrotropic; human; allele.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200290585-A2.  
 XX  
 PD 14-NOV-2002.  
 XX  
 PF 08-MAY-2002; 2002WO-GB02066.  
 XX  
 PR 09-MAY-2001; 2001GB-0011277.  
 XX  
 PA (UYSH-) UNIV SHEFFIELD HALLAM.  
 XX  
 PI El-Nahas AM, Blakemore A, Khalil MS;  
 XX  
 DR WPI; 2003-120560/11.  
 DR N-PSDB; ABV75391.  
 XX  
 PT Determining an individual's susceptibility to the progression of renal  
 PT failure comprises detecting the presence of a genetic polymorphism  
 PT pattern in transforming growth factor beta 1 (TGFβ1) gene in a sample  
 PT from the individual.  
 XX  
 PS Claim 51; Page 59-61; 62pp; English.  
 XX  
 CC The invention relates to determining an individual's susceptibility to  
 CC renal failure and involves detecting the presence of a genetic  
 CC polymorphism pattern in transforming growth factor beta 1 (TGFβ1) gene in  
 CC a sample from an individual, where polymorphism pattern is associated  
 CC with renal failure. The method is useful for determining an individual's  
 CC susceptibility to the progression of renal failure. The nucleic acid  
 CC comprising a T(-509)C polymorphism of TGFβ1 gene, or a polypeptide  
 CC comprising a sequence of 391 amino acids is useful for preparing a  
 CC medicament for retarding or preventing the progression of renal disease,  
 CC and for drug research purposes for retarding or preventing the  
 CC progression of renal disease. Sequences ABV75386-88 represents the  
 CC protein sequence for the TGFβ1 G-allele of the Arg25Pro polymorphism  
 CC of exon 1.

XX SQ Sequence 391 AA;  
 Query Match 88.7%; Score 1916; DB 24; Length 391;  
 Best Local Similarity 89.7%; Pred. No. 2.2e-159;  
 Matches 366; Conservative 10; Mismatches 14; Indels 18; Gaps 2;  
 QY 1 MAPSGRLRLPLLLPDLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
 Db 1 MPPSGRLRLPLLLPDLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
 QY 61 SPFGQVPPGLPEAVLALYNSTRDRVAGESNPEPEADYAKVETRVLMVETHNEI 120  
 Db 61 SPFGQVPPGLPEAVLALYNSTRDRVAGESNPEPEADYAKVETRVLMVETHNEI 120  
 QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRRAELRLRLKLVQEHVELYQKYSNDSW 179  
 Db 121 YDKFKGTSTHSIYMFNTSELREAVPEPVLLSRRAELRLRLKLVQEHVELYQKYSNDSW 180  
 QY 181 YLSNRLAPSDSPWLSFDVTGVVROWLTPREAEIEGFRLSAHSSSDKNTLHVEINGFN 239  
 Db 181 YLSNRLAPSDSPWLSFDVTGVVROWLTPREAEIEGFRLSAHSSSDKNTLHVEINGFN 240  
 QY 241 SGRGDLATIHGMNRPFLLMATPLERAOHLHSSRRRALDNTNSYDVPDVASLALDT 299  
 Db 241 TGRGDLATIHGMNRPFLLMATPLERAOHLHSSRRRALDNTNSYDVPDVASLALDT 283  
 QY 301 YCFSSSTEKNCCVQOLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVIAL 359  
 Db 284 YCFSSSTEKNCCVQOLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVIAL 343  
 QY 361 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 Db 344 NOHNPASAAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 391  
 RESULT 12  
 AAR20124  
 ID AAR20124 standard; Protein: 390 AA.  
 XX  
 AC AAR20124;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 16-APR-1992 (first entry)  
 XX  
 DE Sequence of simian transforming growth factor (TGF) beta-1.  
 XX  
 KW Hypertension therapy; hypotensive agent; blood pressure modulator.  
 XX  
 OS Monkey.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 8..21  
 FT Protein 279..390  
 XX  
 PN WO9119513-A.  
 XX  
 PD 26-DEC-1991.  
 XX  
 PF 20-JUN-1991; 91WO-US04449.  
 XX  
 PR 20-JUN-1990; 90US-0541221.  
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Oleson FB, Comereshki CR;  
 XX  
 DR WPI; 1992-024199/03.  
 DR N-PSDB; AAQ20289.  
 XX  
 PT Use of transforming growth factor (TGF)-beta and their  
 PT antagonists - for modulating blood pressure, for treating  
 PT hypertension and hypotension

```

XX PS Disclosure; Fig 1: 42pp; English.
XX CC
CC CC A new method for treating hypertension comprises administering a
CC transforming growth factor (TGF)-beta to an individual at a dose
CC effective for lowering blood pressure; the TGF-beta may be e.g.
CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-
CC beta2 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-
CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2
CC complex.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX CC
XX CC Sequence 390 AA:
XX CC
XX CC Query Match 88.78; Score 1914.5; DB 13; Length 390;
XX CC Best Local Similarity 89.4%; Pred. No. 2.9e-159;
XX CC Matches 364; Conservative 10; Mismatches 16; Indels 17; Gaps 1;
XX CC
XX CC 1 MAPSGLRLPLLLPLLLVLTGCRPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
XX CC 1 MPPSGLRLPLLLPLLLVLTGCRPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
XX CC
XX CC 61 SPSPQGVPPGPLEPAVLALYNSTRDRVAGESVEPEPEADYAKETVRLMVESGNOI 120
XX CC 61 SPSPQGVPPGPLEPAVLALYNSTRDRVAGESVEPEPEADYAKETVRLMVETHNEI 120
XX CC
XX CC 121 YDKFKGTGPHSLYMLFNTSELREAVPEPVLSSRAELRLRLKLVQEHVELYQKYSNDSWR 180
XX CC 121 YDKFKGTGPHSLYMLFNTSELREAVPEPVLSSRAELRLRLKLVQEHVELYQKYSNDSWR 180
XX CC
XX CC 181 YLSNRLAPSDSPWLSFDVTGVVROWLTRREAIEGFRLSAHSDDSKDNTLHVEINGFN 240
XX CC 181 YLSNRLAPSDSPWLSFDVTGVVROWLTRREAIEGFRLSAHSDDSKDNTLQVDINGFT 240
XX CC
XX CC 241 SGRRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYDPAVASLALDTN 300
XX CC 241 TGRRGDLATIHGMNRPFLMLMATPLERAQHLQSSRRHR-----ALDTN 283
XX CC
XX CC 301 YCFSTSEKNCVRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
XX CC 284 YCFSTSEKNCVRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
XX CC
XX CC 361 NQHNPGASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
XX CC 344 NQHNPGASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390
XX CC
XX CC RESULT 13
XX CC AAR46227
XX CC ID AAR46227 standard; Protein; 390 AA.
XX CC
XX CC AAR46227;
XX CC
XX CC 25-MAR-2003 (updated)
XX CC 09-JUL-1994 (first entry)
XX CC
XX CC Human pre-TGF-beta-1.
XX CC
XX CC TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;
XX CC transforming growth factor beta-3; recombinant; wound healing;
XX CC vulnary.
XX CC
XX CC Homo sapiens.
XX CC
XX CC Key Location/Qualifiers
XX CC Peptide 279..390
XX CC /label- Mat-peptide
XX CC Cleavage-site 279
XX CC /note- "TGF-beta-1 release site"
XX CC Modified-site 82..84
XX CC /label- N-glycosylation_site
XX CC Modified-site 136..138
XX CC /label- N-glycosylation_site

```

```

FT Modified-site 176..178
FT /label- N-glycosylation_site
XX FT
XX PN US284763-A.
XX PD 08-FEB-1994.
XX XX
XX PF 04-MAR-1992; 92US-0845893.
XX XX
XX PR 22-MAR-1985; 85US-0713142.
XX PR 13-MAR-1987; 87US-0025423.
XX PR 04-AUG-1989; 89US-0389929.
XX PR 04-MAR-1992; 92US-0845893.
XX XX
XX (GETH ) GENENTECH INC.
XX PA
XX Derynk RMA, Goeddel DV;
XX PI
XX WPI: 1994-056343/07.
XX DR N-PSDB; AAQ56923.
XX XX
XX Nucleic acid sequences encoding transforming growth factor-beta -
XX diagnostic probes, and for use in therapeutics
XX PS
XX Disclosure; Fig 1b; 25pp; English.
XX CC
XX CC cDNA sequences were determined for human pre-TGF-beta-1 (AAQ56923),
XX CC p19 TGF-beta-3 (AAQ56925) and human TGF-beta-3 (AAQ56926), and the
XX CC corresponding amino acid sequences were determined (AAR46227-29,
XX CC respectively). A genomic fragment corresponding to a human TGF-
XX CC beta-1 exon (AAQ56924) was also isolated and its amino acid sequence
XX CC determined (AAR46230). The sequences have been used in the
XX CC construction of vectors for the expression of recombinant TGF-
XX CC beta.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX CC
XX CC Sequence 390 AA:
XX CC
XX CC Query Match 88.5%; Score 1910.5; DB 15; Length 390;
XX CC Best Local Similarity 89.4%; Pred. No. 6.5e-159;
XX CC Matches 364; Conservative 10; Mismatches 16; Indels 17; Gaps 1;
XX CC
XX CC 1 MAPSGLRLPLLLPLLLVLTGCRPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
XX CC 1 MPPSGLRLPLLLPLLLVLTGCRPAAGLSTCKTIDMELVKRKRIEAIHQILSKRLA 60
XX CC
XX CC 61 SPSPQGVPPGPLEPAVLALYNSTRDRVAGESVEPEPEADYAKETVRLMVESGNOI 120
XX CC 61 SPSPQGVPPGPLEPAVLALYNSTRDRVAGESVEPEPEADYAKETVRLMVETHNEI 120
XX CC
XX CC 121 YDKFKGTGPHSLYMLFNTSELREAVPEPVLSSRAELRLRLKLVQEHVELYQKYSNDSWR 180
XX CC 121 YDKFKGTGPHSLYMLFNTSELREAVPEPVLSSRAELRLRLKLVQEHVELYQKYSNDSWR 180
XX CC
XX CC 181 YLSNRLAPSDSPWLSFDVTGVVROWLTRREAIEGFRLSAHSDDSKDNTLHVEINGFN 240
XX CC 181 YLSNRLAPSDSPWLSFDVTGVVROWLTRREAIEGFRLSAHSDDSKDNTLQVDINGFT 240
XX CC
XX CC 241 SGRRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYDPAVASLALDTN 300
XX CC 241 TGRRGDLATIHGMNRPFLMLMATPLERAQHLQSSRRHR-----ALDTN 283
XX CC
XX CC 301 YCFSTSEKNCVRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
XX CC 284 YCFSTSEKNCVRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
XX CC
XX CC 361 NQHNPGASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
XX CC 344 NQHNPGASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390
XX CC
XX CC RESULT 14
XX CC AAR83054

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 18:00:32 ; Search time 30 Seconds  
(without alignments)  
574.017 Million cell updates/sec

Title: US-10-017-372E-37

Perfect score: 2159

Sequence: 1 MAPSGLRLPLLLPLLLV.....GRKPKVQLSNMIVRSCKCS 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 423:0858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PT05\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1919.5	88.9	390	1 US-08-132-405-1	Sequence 1, Appli
2	1919.5	88.9	390	1 US-08-395-939A-1	Sequence 1, Appli
3	1919.5	88.9	390	5 PCT-US91-01861-1	Patent No. 5168051
4	1919.5	88.9	390	6 5168051-2	Sequence 2, Appli
5	1914.5	88.7	390	1 US-07-669-171-2	Sequence 5, Appli
6	1909	88.4	394	5 PCT-US94-03705-5	Sequence 3, Appli
7	828.5	38.4	412	1 US-08-132-405-3	Sequence 21, Appli
8	828.5	38.4	412	1 US-08-395-939A-3	Sequence 2, Appli
9	824.5	38.2	412	4 US-09-380-662-21	Sequence 2, Appli
10	824.5	38.2	412	5 PCT-US91-04541-2	Sequence 2, Appli
11	824.5	38.2	412	6 5262319-2	Patent No. 5262319
12	823.5	38.1	410	5 PCT-US91-01861-3	Sequence 3, Appli
13	809	37.5	414	1 US-08-132-405-2	Sequence 2, Appli
14	809	37.5	414	1 US-08-395-939A-2	Sequence 6, Appli
15	809	37.5	414	5 PCT-US94-03705-6	Sequence 2, Appli
16	804	37.2	414	5 PCT-US91-01861-2	Sequence 4, Appli
17	803	37.2	304	1 US-08-132-405-4	Sequence 4, Appli
18	803	37.2	304	1 US-08-395-939A-4	Patent No. 5221620
19	803	37.2	304	5 PCT-US91-01861-4	Patent No. 5168051
20	802	37.1	414	6 5221620-4	Patent No. 5221620
21	800.5	37.1	455	6 5168051-11	Sequence 8, Appli
22	793	36.7	442	6 5221620-2	Sequence 1, Appli
23	753.5	34.9	139	4 US-08-065-844A-8	Sequence 36, Appli
24	638	29.6	112	1 US-07-979-441-1	Sequence 41, Appli
25	638	29.6	112	1 US-08-197-792-36	Sequence 41, Appli
26	638	29.6	112	1 US-08-486-057B-41	Sequence 36, Appli
27	638	29.6	112	1 US-08-459-850-36	Sequence 36, Appli

28	638	29.6	112	1 US-08-459-214-36	Sequence 36, Appli
29	638	29.6	112	1 US-08-470-837-30	Sequence 30, Appli
30	638	29.6	112	2 US-08-789-588-41	Sequence 41, Appli
31	638	29.6	112	2 US-08-410-573-1	Sequence 2, Appli
32	638	29.6	112	3 US-09-123-233-2	Sequence 5, Appli
33	638	29.6	112	3 US-08-927-433-5	Sequence 30, Appli
34	638	29.6	112	4 US-08-868-452-30	Sequence 1, Appli
35	638	29.6	112	4 US-09-095-637D-1	Sequence 1, Appli
36	638	29.6	112	5 PCT-US93-03068-1	Sequence 23, Appli
37	638	29.6	114	1 US-08-481-377-23	Sequence 21, Appli
38	638	29.6	114	2 US-08-491-835-21	Sequence 23, Appli
39	638	29.6	114	3 US-09-153-733A-23	Sequence 21, Appli
40	638	29.6	114	3 US-08-946-092A-21	Sequence 21, Appli
41	638	29.6	114	3 US-09-172-062-21	Sequence 22, Appli
42	638	29.6	114	3 US-08-624-635-22	Sequence 21, Appli
43	638	29.6	114	4 US-09-301-520D-21	Sequence 23, Appli
44	638	29.6	114	4 US-09-389-705-23	Sequence 23, Appli
45	638	29.6	114	5 PCT-US94-00666-23	Sequence 23, Appli

## ALIGNMENTS

RESULT 1  
US-08-132-405-1  
; Sequence 1, Application US/08132405  
; Patent No. 5409896  
; GENERAL INFORMATION:  
; APPLICANT: Ammann, Arthur J.  
; TITLE OF INVENTION: Method of Inducing Bone Growth Using  
; TITLE OF INVENTION: TGF-Beta  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/132,405  
; FILING DATE: 06-OCT-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/063841  
; FILING DATE: 18-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/790856  
; FILING DATE: 12-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/401906  
; FILING DATE: 01-SEP-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 597D1C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 390 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-132-405-1





```
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01861-1

Query Match      88.9%; Score 1919.5; DB 5; Length 390;
Best Local Similarity 89.7%; Pred. No. 3.7e-169;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60
Db 1 MPFSGRLRLPLLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60
QY 61 SPPSQGDVPPGPPPLPEAVLALYNSTRDRVAGESVEPEPEADYYAKEVTVRLVMVETNEI 120
Db 61 SPPSQGEVPPGPPPLPEAVLA..YNSTRDRVAGESAEPEPEADYYAKEVTVRLVMVETNEI 120
QY 121 YDKFKGTPTISLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVLYQKYSNDSWR 180
Db 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVLYQKYSNNSWR 180
QY 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVLYQKYSNNSWR 180
Db 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVLYQKYSNNSWR 180
QY 181 YLSNRLLPASDSEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKONTLHVEINGFN 240
Db 181 YLSNRLLPASDSEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKONTLHVEINGFN 240
QY 241 SGRRGDLATIHGMNRPFLLLMATPLERACHLHSSRRRALDTNSYPYDVPDYASLALDTN 300
Db 241 TGRRGDLATIHGMNRPFLLLMATPLERACHLHSSRRRALDTNSYPYDVPDYASLALDTN 300
QY 301 YCFSSTEKNCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLY 360
Db 284 YCFSSTEKNCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLY 343
QY 361 NOHNPASAAAPCCVPQALEPLPIVYVYVGRKPKVEQLSNMIVRSCKCS 407
Db 344 NOHNPASAAAPCCVPQALEPLPIVYVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4
5168051-2
; Patent No. 5168051
; APPLICANT: DERYNCK, RIK M.A.; GORDELL, DAVID V.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TGF-B ITS USES
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/389,929
; FILING DATE: 04-AUG-1989
; SEQ ID NO: 2:
; LENGTH: 390
; 5168051-2

Query Match      88.9%; Score 1919.5; DB 6; Length 390;
Best Local Similarity 89.7%; Pred. No. 3.7e-169;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60
Db 1 MPFSGRLRLPLLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60
QY 61 SPPSQGDVPPGPPPLPEAVLALYNSTRDRVAGESVEPEPEADYYAKEVTVRLVMVETNEI 120
Db 61 SPPSQGEVPPGPPPLPEAVLALYNSTRDRVAGESAEPEPEADYYAKEVTVRLVMVETNEI 120

; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01861-1

Query Match      88.7%; Score 1914.5; DB 1; Length 390;
Best Local Similarity 89.4%; Pred. No. 1.1e-168;
Matches 364; Conservative 10; Mismatches 16; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60
Db 1 MPFSGRLRLPLLLPLLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEATRGQILSKRLA 60
QY 61 SPPSQGDVPPGPPPLPEAVLALYNSTRDRVAGESVEPEPEADYYAKEVTVRLVMVETNEI 120
```

61	SPSOGCEVPPGFLPENVLALNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMEVTHEI	120
121	YDKFKGTPHSYMYLFWTSELSEAVEPVLVLSRAELRLRLKLKVEQHVELYOKYSNDSWR	180
121	YDKFKQSTHSYMFNTSELSEAVEPVLVLSRAELRLRLKLKVEQHVELYOKYSNNSWR	180
181	YLSNRLAPSDSPEWLGFDTVGVVQWMLTRREATEGFRLSAHSSDSKONTLHVETNGFN	240
181	YLSNRLAPLSPSEWLSFDTVGVVQWMLSRGGEIEGFRLSAHSCSDSKONTLQVDINGFT	240
241	SGRRGDLATIGHMNPFFLLWATPLERAQHLHSRRHALDTSYDYPDVPDYASLALDTN	300
241	TGRRGDLATIGHMNPFFLLWATPLERAQHLQSSRRH-----ALDITN	283
301	YCFSSTEKNCCVRLQYIDFRKDLGKWIIHEPKGYHANFCLGPGCPYIWSLDTOYSKVLY	360
284	YCFSSTEKNCCVRLQYIDFRKDLGKWIIHEPKGYHANFCLGPGCPYIWSLDTOYSKVLY	343
361	NOHNPAGAAACCVPQALPLPIVYVYGRKPKVEQLSNMIVRSCKGS	407
344	NOHNPAGAAACCVPQALPLPIVYVYGRKPKVEQLSNMIVRSCKGS	390

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 88.4%; Score 1909; DB 5; Length 394;  
Best Local Similarity 89.5%; Pred. NO. 3.5e-168;

```

Matches 365; Conservative 10; Mismatches 15; Indels 18; Gaps 2
QY 1 MAPSGRLRLLPLLLWLVLTGPRPAAGLSTCKTIIDMFLVKRKRIEAIKQILSKURLA 60
D 1 MPDSGLRLRLLPLLLWLVLTGPPAAGLSTCKTIIDMELVKRKRIEAIKRGQILSKURLA 60
QY 61 SPSSQGDVPGPLPEAVLALYNSTRORVAGESVEPEPEADYYAKEVTVLVMVESGNQI 120
D 61 SPSSQGEVPPGPLPEAVLALYNSTRORVAGESAEPEPEADYYAKEVTIKVLMMVETHEI 120
QY 121 YDFKCTPHSLYMLFNTSELREAVPVPVLLSRAELRL--RLKI.KVEGHVELYKYYSNDSW 179
D 121 YDAFKQSTHSIYWFNTSELREAVPEPVLLSRAELRLRLKLKVEQHVELYKYYSNNWSW 180
QY 180 RYLNSNLLAPSDPENLSFDVTGVVQWLITRRRAIEGFRLSAHSSSDSKONTLHVEINGF 239
D 181 RYLNSNLLAPSDPENLSFDVTGVVQWLISRGGELECFRLSAHCSCDSRDNTIQVDINGF 240
QY 240 NSGRRGDLATIHGMNRPFLLMATPLERAQHLHSRRHRALDINSYPDYDPDVASLALDT 299
D 241 TTGRRGDLATIHGMNRPFLLMATPLERAQHLQSSRHRR-----ALDT 283
QY 300 NYCSFSSTEKNCCVRQLYIDFRKDOLGWKHKEHPKGYHANFCLGPCPYTWSLDTOYSKVIAL 359
D 284 NYCSFSSTEKNCCVRQLYIDFRKDOLGWKHKEHPKGYHANFCLGPCPYTWSLDTOYSKVIAL 343
QY 360 YNOHNFGASAAPCCVQALESPLPIVVYVGPKPVEQLSNMIVRSCKKS 407
D 344 YNOHNFGASAAPCCVQALESPLPIVVYVGPKPVEQLSNMIVRSCKKS 391

RESULT 7
US-08-132-405-3
: Sequence 3, Application US/08132405
: Patent No. 5409896
: GENERAL INFORMATION:
: APPLICANT: Ammann, Arthur J.
: APPLICANT: Rudman, Christopher G.
: TITLE OF INVENTION: Method of Inducing Bone Growth Using
: TITLE OF INVENTION: TGF-Beta
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/132,405
: FILING DATE: 06-OCT-1993
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/063841
: FILING DATE: 18-MAY-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/790856
: FILING DATE: 12-NOV-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/401906
: FILING DATE: 01-SEP-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Hasak, Janet E.
: REGISTRATION NUMBER: 28,616
: REFERENCE/DOCKET NUMBER: 597DLC2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1896
: TELEFAX: 415/952-9881

```

```

; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-132-405-3

Query Match      38.4%; Score 828.5; DB 1; Length 412;
Best Local Similarity 44.5%; Pred. No. 2.6e-68;
Matches 189; Conservative 57; Mismatches 126; Indels 53; Gaps 13;

QY 15 LMLLVLTGPRPAAGLSTCTIDMELVKRKRIEARGQILSKRLASPPSQGDVPPGPLP 74
Db 9 LVVALLNATVSVLSLSTCTIDFGHIKKRVEAIRGQILSKRLTSPPEPTVMT--HVP 66

QY 75 EAVLALYNSTRORVA--GESVE---PEPEPEADYAKVTRVLMVE---SGNQIYDKFKG 126
Db 67 YQVVALYNSTRRELLEHGERKEEGCTQENTSEYAKETHKFDMIQGLAEHNELAVCPKG 126

QY 127 TPHSLYALNTSELRAVPPEVLLSRAELRLRL---KLKVEQHVLYQKYSND----S 178
Db 127 ITSKVFR-FNVSSVEK---NRTNLFRAEFRLVVPNPSSKRNEQRIELFQILRPDEHIAK 182

QY 179 WRYLSNRLAPSDSPWLSFDVTGVVRQWLTRREALTEGFRLSAH-----SSDSKDN- 230
Db 183 QRYIGCKNLPTRGTAEWLSFDVDTVREWLLRESNLGLEISIHCPCHTFQPNGDILENI 242

QY 231 --TLHVEINGFNS---GRRGDLATIGHM---NRPFLLLMATPLERAQHLHSSRRALDT 282
Db 243 HEYMEIKFKGVNEDDHGRGDLGLRLLKQKDNPNPHILMMIPPHRLDNPQGGQGRK--- 298

QY 283 NSPYDVPDYASLALDNTNYCFSTEKNCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGP 342
Db 343 CPYIWSLDTQYSKVLALYNQHNPGASAPCCVQALEPLPIVYVGRKPKVEQLSNMIVR 402

QY 348 CPYLSADTTHTSVGLYNTLNPEASASPCCVQDLEPLTILYVGRTPKVEQLSNMIVK 407
Db 403 SCCKS 407
Db 408 SCCKS 412

RESULT 8
US-08-395-939A-3
; Sequence 3, Application US/08395939A
; Patent No. 5604204
; GENERAL INFORMATION:
; APPLICANT: Ammann, Arthur J.
; APPLICANT: Rudman, Christopher G.
; TITLE OF INVENTION: TGF-BETA COMPOSITION FOR INDUCING BONE
; TITLE OF INVENTION: GROWTH
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 350 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,939A
; FILING DATE: 27-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/132405

```

```

; FILING DATE: 12-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/063841
; FILING DATE: 18-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790856
; FILING DATE: 12-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/401906
; FILING DATE: 1-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28 616
; REFERENCE/DOCKET NUMBER: P0597D1C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-395-939A-3

Query Match      38.4%; Score 828.5; DB 1; Length 412;
Best Local Similarity 44.5%; Pred. No. 2.6e-68;
Matches 189; Conservative 57; Mismatches 126; Indels 53; Gaps 13;

QY 15 LMLLVLTGPRPAAGLSTCTIDMELVKRKRIEARGQILSKRLASPPSQGDVPPGPLP 74
Db 9 LVVALLNATVSVLSLSTCTIDFGHIKKRVEAIRGQILSKRLTSPPEPTVMT--HVP 66

QY 75 EAVLALYNSTRORVA--GESVE---PEPEPEADYAKVTRVLMVE---SGNQIYDKFKG 126
Db 67 YQVVALYNSTRRELLEHGERKEEGCTQENTSEYAKETHKFDMIQGLAEHNELAVCPKG 126

QY 127 TPHSLYALNTSELRAVPPEVLLSRAELRLRL---KLKVEQHVLYQKYSND----S 178
Db 127 ITSKVFR-FNVSSVEK---NRTNLFRAEFRLVVPNPSSKRNEQRIELFQILRPDEHIAK 182

QY 179 WRYLSNRLAPSDSPWLSFDVTGVVRQWLTRREALTEGFRLSAH-----SSDSKDN- 230
Db 183 QRYIGCKNLPTRGTAEWLSFDVDTVREWLLRESNLGLEISIHCPCHTFQPNGDILENI 242

QY 231 --TLHVEINGFNS---GRRGDLATIGHM---NRPFLLLMATPLERAQHLHSSRRALDT 282
Db 243 HEYMEIKFKGVNEDDHGRGDLGLRLLKQKDNPNPHILMMIPPHRLDNPQGGQGRK--- 298

QY 283 NSPYDVPDYASLALDNTNYCFSTEKNCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGP 342
Db 299 -----KRALDNTNYCFRNLENCVCRPLYIDFRQDLGKWKVHEPKGYANFCSGP 347

QY 343 CPYIWSLDTQYSKVLALYNQHNPGASAPCCVQALEPLPIVYVGRKPKVEQLSNMIVR 402
Db 348 CPYLSADTTHTSVGLYNTLNPEASASPCCVQDLEPLTILYVGRTPKVEQLSNMIVK 407

QY 403 SCCKS 407
Db 408 SCCKS 412

RESULT 9
US-09-380-662-21
; Sequence 21, Application US/09380662
; Patent No. 6376199
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF TROPHOBLAST
; FILE REFERENCE: 11757.39USNO
; CURRENT APPLICATION NUMBER: US/09/380,662

```

```
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent version 3.0
; SEQ ID NO 21
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-380-662-21

Query Match      38.2%; Score 824.5; DB 4; Length 412;
Best Local Similarity 44.2%; Pred. No. 6.1e-68;
Matches 188; Conservative 58; Mismatches 126; Indels 53; Gaps 13;

QY 15 LMLVLTPGRPAAGLSTCKTIDMELVKRRIEAIKQILSKRLASPPSGDVPPLP 74
Db 9 LVVLLNFATVSLSTCTTLDGHIKKRVEAIRGQILSKRLTSPPEPTVMT--HVP 66

QY 75 EAVLYNSTR---DRVAGESVE--PEPEADYYAKEVTRVLMVE---SGNQIYDKFKG 126
Db 67 QVLYALYNSTRLEEMHGEREGCTQENTSEYAKEIHKEDMIOGLAEHNLAVCPKG 126

QY 127 TPHSLYMLNTPSELRAVPEPVLLSRAELRLRL---KLKVEQHVLYQKYSND----S 178
Db 127 ITSQVFR-FNVSSVEK---NRTNLFRAEFVRLVVPNPSKNEQRIELFQILRDEHIAK 182

QY 179 WYLSNRLAPSDSPWLSFQVTVGVVQWLTTRAELEGFRLSAH-----SSDSKDN- 230
Db 183 QVIGGKNLPTGTAEWLSFQVTVGVVQWLTTRAELEGFRLSAH-----SSDSKDN- 242

QY 231 --TLHVEINGFNS---GRRGLATIGHM---NRPLLMLATPLERAQHLHSSRRALDT 282
Db 243 HEVMEIKFGVDNEDDHGRDLGRLLKQKHNNPHLLMNPPLHNDNPGQGGQK---- 298

QY 283 NSYPYDVPDYASIALDTNYCFSSTEKNCVQRYLYIDFRKDLGKWIHPKGYHANFCLGP 342
Db 299 -----KRALDTNYCFRNLEENCVRPLYIDFRQDLGKWWHPKGYANFCSGP 347

QY 343 CPYIWSLDTQYKVLALYNQHNPGASAPCCVPQALEPLPIVYVYGRKPKVEQLSNMIVR 402
Db 348 CPYLSADTTHSTVGLYNTLNPEASAPCCVPQDLEPLTILYVYGRTPKVEQLSNMIVK 407

QY 403 SKCS 407
Db 408 SKCS 412

RESULT 10
PCT-US91-04541-2
; Sequence 2, Application PC/TUS9104541
; GENERAL INFORMATION:
; APPLICANT: Oncogene Science Inc.
; TITLE OF INVENTION: TISSUE DERIVED TUMOR GROWTH INHIBITORS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/04541
; FILING DATE: 19910625
; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1919/22669-F-PCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523 COOP U1
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US91-04541-2

Query Match      38.2%; Score 824.5; DB 5; Length 412;
Best Local Similarity 44.2%; Pred. No. 6.1e-68;
Matches 188; Conservative 58; Mismatches 126; Indels 53; Gaps 13;

QY 15 LMLVLTPGRPAAGLSTCKTIDMELVKRRIEAIKQILSKRLASPPSGDVPPLP 74
Db 9 LVVLLNFATVSLSTCTTLDGHIKKRVEAIRGQILSKRLTSPPEPTVMT--HVP 66

QY 75 EAVLYNSTR---DRVAGESVE--PEPEADYYAKEVTRVLMVE---SGNQIYDKFKG 126
Db 67 QVLYALYNSTRLEEMHGEREGCTQENTSEYAKEIHKEDMIOGLAEHNLAVCPKG 126

QY 127 TPHSLYMLNTPSELRAVPEPVLLSRAELRLRL---KLKVEQHVLYQKYSND----S 178
Db 127 ITSQVFR-FNVSSVEK---NRTNLFRAEFVRLVVPNPSKNEQRIELFQILRDEHIAK 182

QY 179 WYLSNRLAPSDSPWLSFQVTVGVVQWLTTRAELEGFRLSAH-----SSDSKDN- 230
Db 183 QVIGGKNLPTGTAEWLSFQVTVGVVQWLTTRAELEGFRLSAH-----SSDSKDN- 242

QY 231 --TLHVEINGFNS---GRRGLATIGHM---NRPLLMLATPLERAQHLHSSRRALDT 282
Db 243 HEVMEIKFGVDNEDDHGRDLGRLLKQKHNNPHLLMNPPLHNDNPGQGGQK---- 298

QY 283 NSYPYDVPDYASIALDTNYCFSSTEKNCVQRYLYIDFRKDLGKWIHPKGYHANFCLGP 342
Db 299 -----KRALDTNYCFRNLEENCVRPLYIDFRQDLGKWWHPKGYANFCSGP 347

QY 343 CPYIWSLDTQYKVLALYNQHNPGASAPCCVPQALEPLPIVYVYGRKPKVEQLSNMIVR 402
Db 348 CPYLSADTTHSTVGLYNTLNPEASAPCCVPQDLEPLTILYVYGRTPKVEQLSNMIVK 407

QY 403 SKCS 407
Db 408 SKCS 412

RESULT 11
5262319-2
; Patent No. 5262319
; APPLICANT: IWATA, KENNETH K.; FOULKES, J. GORDON; DIJKE,
; PETER T.; HALEY, JOHN D.
; TITLE OF INVENTION: METHOD FOR OBTAINING BONE MARROW FREE
; OF TUMOR CELLS USING TRANSFORMING GROWTH FACTOR R3
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/543,341
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 353,410
; FILING DATE: 17-MAY-1989
; APPLICATION NUMBER: 183,410
; FILING DATE: 20-APR-1988
; APPLICATION NUMBER: 111,022
; FILING DATE: 20-OCT-1987
; APPLICATION NUMBER: 922,121
; FILING DATE: 20-OCT-1986
```

```

; APPLICATION NUMBER: 847,931
; FILING DATE: 07-APR-1986
; APPLICATION NUMBER: 725,003
; FILING DATE: 19-APR-1985
; SEQ ID NO:2
; LENGTH: 412
5262319-2

```

```

Query Match      38.2%; Score 824.5; DB 6; Length 412;
Best Local Similarity 44.2%; Pred. No. 6.1e-68;
Matches 188; Conservative 58; Mismatches 126; Indels 53; Gaps 13;

Qy 15 LMLLVLTGCPAAGLSTCKTIDMELVKRKRIRAEIRGOILSKRLASPPSGQDVPPGPLP 74
Db 9 LVVALLNFATVLSLSSTCTTDFGHKKRKRVEAIRGOILSKRLTSPPEPTVMT--HVP 66
Qy 75 EAVLALYNSTR--DRVAGESVE--PEPEADYAKERVRLMVE---SGNOIYDKFKG 126
Db 67 YQVLALYNSTRLEENHGEREGCTOENTSEYAYAKEIHKFDMIOGLAEHNE LAVCPKG 126
Qy 127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLRL---KLKVEQHVELYQKYSND---S 178
Db 127 ITSQVFR-FNVSSVEK---NRTNLFRAEFRLVVPNPSSKRNEQRIELFOILRPDEHIAK 182
Qy 179 WRYLSNRLAPSDPEWLSFDVTGVVROWLTRREAIEGFRLSAH-----SSSDSKDN- 230
Db 183 QRYIGGKNLPTRGTAEWLSFDVDTVREMLLRRESNLGLEISIHCPCHTFQPNGDILENI 242
Qy 231 --TLHVEINGNS---GRGDLATIHGM---NRPFLLMATPLERAQHLHSSRRRALDT 282
Db 243 HEVMEIKFGVDNEDHGRGDLGRLLKQKQDHHNPHLILMIPPHRLDNPQGQGRK---- 298
Qy 283 NSYPVDYASLALDYNCFSSSTEKNCCVRQLYIDFKDLGKWKIHEPKGYHANFCLGP 342
Db 299 -----KRALDINFCRNLECCVRRPLYIDFQDLGKWKIHEPKGYHANFCSGP 347
Qy 343 CPYIWSLDTQYSKVLALYNQHNFCASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVR 402
Db 348 CPYLSADTTHSTVLGLYNTLNPEASAPCCVPODLEPLTILYVVGRTPKVEQLSNMVK 407
Qy 403 SKCS 407
Db 408 SKCS 412

```

```

RESULT 12
PCT-US91-01861-3
; Sequence 3, Application PC/TUS9101861
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Method of Predispersing Mammals to
; TITLE OF INVENTION: Accelerated Tissue Repair
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01861
; FILING DATE: 19910320
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Ser. No. 07/504,495
; FILING DATE: 4 April 1990
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US91-01861-3

Query Match      38.1%; Score 823.5; DB 5; Length 410;
Best Local Similarity 44.2%; Pred. No. 7.5e-68;
Matches 188; Conservative 58; Mismatches 126; Indels 53; Gaps 13;

Qy 15 LMLLVLTGCPAAGLSTCKTIDMELVKRKRIRAEIRGOILSKRLASPPSGQDVPPGPLP 74
Db 7 LVVALLNFATVLSLSSTCTTDFGHKKRKRVEAIRGOILSKRLTSPPEPTVMT--HVP 64
Qy 75 EAVLALYNSTRDVA--GESVE--PEPEADYAKERVRLMVE---SGNOIYDKFKG 126
Db 65 YQVLALYNSTRLEENHGEREGCTOENTSEYAYAKEIHKFDMIOGLAEHNE LAVCPKG 124
Qy 127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLRL---KLKVEQHVELYQKYSND---S 178
Db 125 ITSQVFR-FNVSSVEK---NRTNLFRAEFRLVVPNPSSKRNEQRIELFOILRPDEHIAK 180
Qy 179 WRYLSNRLAPSDPEWLSFDVTGVVROWLTRREAIEGFRLSAH-----SSSDSKDN- 230
Db 181 QRYIGGKNLPTRGTAEWLSFDVDTVREMLLRRESNLGLEISIHCPCHTFQPNGDILENI 240
Qy 231 --TLHVEINGNS---GRGDLATIHGM---NRPFLLMATPLERAQHLHSSRRRALDT 282
Db 241 HEVMEIKFGVDNEDHGRGDLGRLLKQKQDHHNPHLILMIPPHRLDNPQGQGRK---- 296
Qy 283 NSYPVDYASLALDYNCFSSSTEKNCCVRQLYIDFKDLGKWKIHEPKGYHANFCLGP 342
Db 297 -----KRALDINFCRNLECCVRRPLYIDFQDLGKWKIHEPKGYHANFCSGP 345
Qy 343 CPYIWSLDTQYSKVLALYNQHNFCASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVR 402
Db 346 CPYLSADTTHSTVLGLYNTLNPEASAPCCVPODLEPLTILYVVGRTPKVEQLSNMVK 405
Qy 403 SKCS 407
Db 406 SKCS 410

RESULT 13
US-08-132-405-2
; Sequence 2, Application US/08132405
; Patent No. 5409896
; GENERAL INFORMATION:
; APPLICANT: Ammann, Arthur J.
; APPLICANT: Rudman, Christopher G.
; TITLE OF INVENTION: Method of Inducing Bone Growth Using
; TITLE OF INVENTION: TGF-Beta
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/132,405
; FILING DATE: 06-OCT-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/063841
; FILING DATE: 18-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790856
; FILING DATE: 12-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/401906
; FILING DATE: 01-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 597D1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-132-405-2

Query Match 37.5%; Score 809; DB 1; Length 414;
Best Local Similarity 43.2%; Pred. No. 1.7e-66;
Matches 181; Conservative 62; Mismatches 116; Indels 60; Gaps 14;

QY 27 AAGLSTCKTIDMELVKRRKRIEARGQILSKRLASPPSQGDVP-PGPLPAVLALYNSTR 85
DB 18 ALSLSTCTLDMDQFMRRKRIEARGQILSKRLKLTSPPE--DYPEPEVPEVISIYNSTR 75

QY 86 D-----RVAGESVEPEPE-PEADYIAKEVTRVLM---VESGNQIYDFKGTGPHSLYMLFNT 137
DB 76 DLLQEKASRRAAACERERSDEEYIAKEVYKIDMPFFPFSNAIPPTFY-RPYFRIVRFDV 134

QY 138 SELREAVPEPVLSSRAELRLRL---KLKV-EQHVELYQ-----KYSNDSWRVLSNRLLA 188
DB 135 SAMEKNASN---LVKAEFRVRLQNKARVPEQRIELYQLKSKDLTSPRTQYIDSKVVK 191

QY 189 PSDSPEWLSFDVTGVVROWLTRREAIFGFRLSAHSSS-----DSKDNTHLVE-- 235
DB 192 TRAGEWLSFDVTDVAVHEWHLHKDRNLGFKLSLHCPCTFTVPSNNYIIPNKSELEARFA 251

QY 236 -INGFNSGRRGLDATHGMNR-----PFLLLMATPLERAQHLHSSRRRALDNTNSPYD 288
DB 252 GIDGTSTVTSQDKTIKSTRKKNKGKTPHULLMLLPSYRLSESQOTNRRKKR----- 302

QY 289 VPDYASLALDNTNYCFSTEKNCCVRLYIDFKDLGKWKIHEPKGYHANFCLGPGCPYIWS 348
DB 303 -----ALDAAACYFRVQNCCLRLYIDFKDLGKWKIHEPKGYNANFCAGACPYLWS 355

QY 349 LDQYSKVLALYNQHNPCASAPCCVPOALEPLPIVYVVGKPKVQGLSNMIVRSCKCS 407
DB 356 SDTQHSRVLSLYNTINPEASASPCCVSQDLEPLTILYIYIGKTPKIEQLSNMIVRSCKCS 414

RESULT 14
US-08-395-939A-2
; Sequence 2, Application US/08394939A
; Patent No. 5604204
; GENERAL INFORMATION:
; APPLICANT: Ammann, Arthur J.
; APPLICANT: Rudman, Christopher G.
; TITLE OF INVENTION: TGF-BETA COMPOSITION FOR INDUCING BONE
; TITLE OF INVENTION: GROWTH
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

```

```

; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,939A
; FILING DATE: 27-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/132405
; FILING DATE: 12-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/063841
; FILING DATE: 18-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790856
; FILING DATE: 12-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/401906
; FILING DATE: 1-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: P0597D1C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-395-939A-2

Query Match 37.5%; Score 809; DB 1; Length 414;
Best Local Similarity 43.2%; Pred. No. 1.7e-66;
Matches 181; Conservative 62; Mismatches 116; Indels 60; Gaps 14;

QY 27 AAGLSTCKTIDMELVKRRKRIEARGQILSKRLASPPSQGDVP-PGPLPAVLALYNSTR 85
DB 18 ALSLSTCTLDMDQFMRRKRIEARGQILSKRLKLTSPPE--DYPEPEVPEVISIYNSTR 75

QY 86 D-----RVAGESVEPEPE-PEADYIAKEVTRVLM---VESGNQIYDFKGTGPHSLYMLFNT 137
DB 76 DLLQEKASRRAAACERERSDEEYIAKEVYKIDMPFFPFSNAIPPTFY-RPYFRIVRFDV 134

QY 138 SELREAVPEPVLSSRAELRLRL---KLKV-EQHVELYQ-----KYSNDSWRVLSNRLLA 188
DB 135 SAMEKNASN---LVKAEFRVRLQNKARVPEQRIELYQLKSKDLTSPRTQYIDSKVVK 191

QY 189 PSDSPEWLSFDVTGVVROWLTRREAIFGFRLSAHSSS-----DSKDNTHLVE-- 235
DB 192 TRAGEWLSFDVTDVAVHEWHLHKDRNLGFKLSLHCPCTFTVPSNNYIIPNKSELEARFA 251

QY 236 -INGFNSGRRGLDATHGMNR-----PFLLLMATPLERAQHLHSSRRRALDNTNSPYD 288
DB 252 GIDGTSTVTSQDKTIKSTRKKNKGKTPHULLMLLPSYRLSESQOTNRRKKR----- 302

QY 289 VPDYASLALDNTNYCFSTEKNCCVRLYIDFKDLGKWKIHEPKGYHANFCLGPGCPYIWS 348
DB 303 -----ALDAAACYFRVQNCCLRLYIDFKDLGKWKIHEPKGYNANFCAGACPYLWS 355

QY 349 LDQYSKVLALYNQHNPCASAPCCVPOALEPLPIVYVVGKPKVQGLSNMIVRSCKCS 407
DB 356 SDTQHSRVLSLYNTINPEASASPCCVSQDLEPLTILYIYIGKTPKIEQLSNMIVRSCKCS 414

```

RESULT 15

PCT-US94-03705-6  
 ; Sequence 6, Application PC/TUS9403705  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mu-Ed Lee  
 ; APPLICANT: Mark A. Perrella  
 ; TITLE OF INVENTION: TRANSFORMING GROWTH  
 ; TITLE OF INVENTION: FACTOR- INHIBITS  
 ; TITLE OF INVENTION: INDUCIBLE NITRIC OXIDE  
 ; TITLE OF INVENTION: SYNTHASE GENE  
 ; TITLE OF INVENTION: TRANSCRIPTION  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM PS/2 Model 50z or 55SX  
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)  
 ; SOFTWARE: WordPerfect (Version 5.1)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/03705  
 ; FILING DATE: 5 April 1994  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Janis K. Fraser  
 ; REGISTRATION NUMBER: Reg. No. 34,819  
 ; REFERENCE/DOCKET NUMBER: 05433/007001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-507C  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 414  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; PCT-US94-03705-6

Query Match 37.5%; Score 809; DB 5; Length 414;  
 Best Local Similarity 43.2%; Pred. No. 1.7e-66;  
 Matches 181; Conservative 62; Mismatches 116; Indels 60; Gaps 14;

QY	27	AAGLSTCKTIDMELVKRKEIAIRGQILSKRLASPPSQGDVP-PGPLPEAVIALYNSTR	85
DB	18	ALSLSTCTLDQFMRRKRIEIRGQILSKRLKLTSPPE--DYPEEVPPEVVISIYNSTR	75
QY	86	D---RVAGESVEPEPE-EEADYAKAEVTRVLM---VESGNIYDKFKGTPHSLYMLFNT	137
DB	76	DLQEKASRRRAACERSEDEEYAKEVYKIDMPFPFSENAIPPTFY-RPYFRIVRFDV	134
QY	138	SELREAVPEPVLLSRAELFLRL---KLKV-EQHVELYQ-----KYSNDSWRYLSNLLA	188
DB	135	SAMEKNASN--LVKAEPFVFRQNPKARVPEQRIELYQILKSKDLTSPTQRYIDSKVKV	191
QY	189	PSDSEWLSFDVTGVVROWLTRRAIEGFRLSAHSSS-----DSKNTLHVE--	235
DB	192	TRAEGEWLSFDVTDAVHEWLLHKORNLGFKISLHCPCTFVPSNNYIIPNKSELEAREFA	251
QY	236	INGFNSGRRGDLATIHGNNR-----PFLIMATPLERAOHLASSRRHRALDTSNYPYD	288
DB	252	GIDGTSTYSGDQKTIKSTRKKNCKTPHLLMLLPYSRLESQQTNRKKR-----	302
QY	289	VPDYASALDNTNYCFSSFEKNCCVRQLYIDFRKDLGNKWIHEPKGYHANFCLGFCPYIWS	348

Db	303	-----ALDAAYCFRNVQDNCCLRLPLYIDFRKDLGNKWIHEPKGYNANFCAGACPYLWS	355
QY	349	LDTQYSKVLALYNQHNPGASAAAPCCVPALEPLPIVYVYVGRKPKVQOLSNMIVRSCKCS	407
DB	356	SDTQHSRVLSLYNTINPEASASPCCVSQDLEPLTILYIYIGKTPKIEOLSNMIVKSKCS	414

Search completed: October 7, 2003, 18:05:41  
 Job time : 31 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 18:04:22 ; Search time 67 Seconds  
(without alignments)  
961.085 Million cell updates/sec

Title: US-10-017-372E-37  
Perfect score: 2159  
Sequence: 1 MAPSGLRLPLLLPLMLLV.....GRKPRVQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 15812981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2030.5	94.0	390	11 US-09-214-592-26	Sequence 26, Appl
2	1929.5	89.4	390	11 US-09-214-592-29	Sequence 29, Appl
3	1926.5	89.2	390	15 US-10-087-268-2	Sequence 2, Appl
4	1919.5	88.9	390	15 US-10-087-268-5	Sequence 5, Appl
5	1916.5	88.8	390	12 US-10-276-947-1	Sequence 1, Appl
6	1914.5	88.7	390	11 US-09-214-592-33	Sequence 33, Appl
7	1909	88.4	391	11 US-09-214-592-17	Sequence 17, Appl
8	1905.5	88.3	390	10 US-09-756-283A-23	Sequence 23, Appl
9	1893.5	87.7	390	11 US-09-214-592-28	Sequence 28, Appl
10	1835.5	85.0	390	11 US-09-214-592-20	Sequence 20, Appl
11	1835.5	85.0	390	11 US-09-214-592-23	Sequence 23, Appl
12	1572.5	72.8	315	11 US-09-214-592-25	Sequence 25, Appl
13	1262.5	58.5	455	10 US-09-756-283A-20	Sequence 20, Appl
14	1139	52.8	447	10 US-09-756-283A-22	Sequence 22, Appl
15	1047	48.5	373	11 US-09-214-592-32	Sequence 32, Appl

16	871.5	40.4	412	11	US-09-214-592-31	Sequence 31, Appl
17	868.5	40.2	382	11	US-09-214-592-34	Sequence 34, Appl
18	868	40.2	383	10	US-09-756-283A-27	Sequence 27, Appl
19	834	38.6	409	11	US-09-214-592-27	Sequence 27, Appl
20	827.5	38.3	410	11	US-09-214-592-22	Sequence 22, Appl
21	826.5	38.3	412	11	US-09-214-592-24	Sequence 24, Appl
22	824.5	38.2	412	11	US-09-214-592-19	Sequence 19, Appl
23	824.5	38.2	412	14	US-10-028-158-21	Sequence 21, Appl
24	824	38.2	414	11	US-09-214-592-21	Sequence 21, Appl
25	820.5	38.0	412	10	US-09-756-283A-25	Sequence 25, Appl
26	809	37.5	414	10	US-09-756-283A-24	Sequence 24, Appl
27	809	37.5	414	11	US-09-214-592-18	Sequence 18, Appl
28	808	37.4	412	11	US-09-214-592-30	Sequence 30, Appl
29	792	36.7	304	10	US-09-756-283A-26	Sequence 26, Appl
30	753.5	34.9	139	14	US-10-002-278-8	Sequence 8, Appl
31	638	29.6	112	10	US-09-813-398-13	Sequence 13, Appl
32	638	29.6	113	10	US-09-813-398-13	Sequence 13, Appl
33	638	29.6	114	10	US-09-813-459-22	Sequence 22, Appl
34	638	29.6	114	14	US-10-115-406-21	Sequence 21, Appl
35	638	29.6	114	15	US-10-154-333-23	Sequence 23, Appl
36	638	29.6	115	10	US-09-859-211-47	Sequence 47, Appl
37	638	29.6	115	10	US-09-880-708-25	Sequence 25, Appl
38	638	29.6	115	11	US-09-872-856-47	Sequence 47, Appl
39	638	29.6	115	15	US-10-335-483-29	Sequence 29, Appl
40	562	26.0	98	12	US-10-187-394-1	Sequence 1, Appl
41	542	25.1	116	14	US-10-115-406-24	Sequence 24, Appl
42	542	25.1	116	15	US-10-154-333-26	Sequence 26, Appl
43	535	24.9	112	10	US-09-813-271B-8	Sequence 8, Appl
44	504	23.3	112	10	US-09-813-271B-12	Sequence 12, Appl
45	504	23.3	114	14	US-10-115-406-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1  
US-09-214-592-26  
; Sequence 26, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki,CMotoo  
; APPLICANT: Shibata,Ckenjl  
; APPLICANT: Sato,Cyasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 26  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: porcine  
; US-09-214-592-26

Query Match	94.0%	Score	2030.5	DB	11	Length	390
Best Local Similarity	94.8%	Pred. No.	3.9e+180				
Matches	386	Conservative	1	Mismatches	3	Indels	17
Gaps	1						
Qy	1	MAPSGURLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIIRGQILSKRLA	60				
Db	1	MPPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIIRGQILSKRLA	60				
Qy	61	SPFSQGDVPPGPPPEAVLALYNSTRVAGESVEPEPEADYIAKEVTRVLMVESGNOI	120				
Db	61	SPFSQGDVPPGPPPEAVLALYNSTRVAGESVEPEPEADYIAKEVTRVLMVESGNOI	120				
Qy	121	YDKFKGTGPHSLYMLFNTSELREAVPEVLLSRALRLRLKLVKVEOHVLYQKYSNDSWR	180				
Db	121	YDKFKGTGPHSLYMLFNTSELREAVPEVLLSRALRLRLKLVKVEOHVLYQKYSNDSWR	180				
Qy	181	YLSNRLAPSDSPWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKNTLHVEINGFN	240				

```

Db      181 YLSNRLAPSDPEWLSFDVTGVVQWLTTRRAIEGFRLSAHCSDSKDNTLHVEINGFN 240
QY      241 SGRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYVASYALDNTN 300
Db      241 SGRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYVASYALDNTN 283
QY      301 YCFSTKNCVQRLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
Db      284 YCFSTKNCVQRLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
QY      361 NOHNPASAPCCVQPALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
Db      344 NOHNPASAPCCVQPALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 2
US-09-214-592-29
; Sequence 29, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki, Cmotoo
; APPLICANT: Shibata, Ckenji
; APPLICANT: Sato, Cyasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 29
; LENGTH: 390
; TYPE: PRT
; ORGANISM: ovine
US-09-214-592-29

Query Match      89.4%; Score 1929.5; DB 11; Length 390;
Best Local Similarity 89.7%; Pred. No. 9.4e-171;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY      1 MAPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKRLA 60
Db      1 MPPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKRLA 60
QY      61 SPPSQGDVPPGPLEAVLALYNSTRDVAGESVEPEPEADYAKETVRLMWESGNOI 120
Db      61 SPPSQGDVPPGPLEAVLALYNSTRDVAGESAETPEPEADYAKETVRLMWETGNI 120
QY      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVQHVLYQKYSNDSWR 180
Db      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVQHVLYQKYSNDSWR 180
QY      181 YLSNRLAPSDPEWLSFDVTGVVQWLTTRRAIEGFRLSAHCSDSKDNTLHVEINGFN 240
Db      181 YLSNRLAPSDPEWLSFDVTGVVQWLTTRRAIEGFRLSAHCSDSKDNTLHVEINGFN 240
QY      241 SGRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYVASYALDNTN 300
Db      241 SGRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYVASYALDNTN 283
QY      301 YCFSTKNCVQRLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
Db      284 YCFSTKNCVQRLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
QY      361 NOHNPASAPCCVQPALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
Db      344 NOHNPASAPCCVQPALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 3
US-10-087-268-2
; Sequence 2, Application US/10087268
; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsson, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condit
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-10-087-268-2

Query Match      88.9%; Score 1919.5; DB 15; Length 390;
Best Local Similarity 89.7%; Pred. No. 8e-170;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY      1 MAPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKRLA 60
```

```

; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsson, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condit
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-10-087-268-2

Query Match      89.2%; Score 1926.5; DB 15; Length 390;
Best Local Similarity 89.9%; Pred. No. 1.8e-170;
Matches 366; Conservative 10; Mismatches 14; Indels 17; Gaps 1;

QY      1 MAPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKRLA 60
Db      1 MPPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKRLA 60
QY      61 SPPSQGDVPPGPLEAVLALYNSTRDVAGESVEPEPEADYAKETVRLMWESGNOI 120
Db      61 SPPSQGDVPPGPLEAVLALYNSTRDVAGESAETPEPEADYAKETVRLMWETGNI 120
QY      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVQHVLYQKYSNDSWR 180
Db      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVQHVLYQKYSNDSWR 180
QY      181 YLSNRLAPSDPEWLSFDVTGVVQWLTTRRAIEGFRLSAHCSDSKDNTLHVEINGFN 240
Db      181 YLSNRLAPSDPEWLSFDVTGVVQWLTTRRAIEGFRLSAHCSDSKDNTLHVEINGFN 240
QY      241 SGRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYVASYALDNTN 300
Db      241 SGRGDLATIHGMNRPFLMLMATPLERAQHLHSSRRRALDNTNSYPDYVASYALDNTN 283
QY      301 YCFSTKNCVQRLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
Db      284 YCFSTKNCVQRLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
QY      361 NOHNPASAPCCVQPALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
Db      344 NOHNPASAPCCVQPALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4
US-10-087-268-5
; Sequence 5, Application US/10087268
; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsson, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condit
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-10-087-268-5

Query Match      88.9%; Score 1919.5; DB 15; Length 390;
Best Local Similarity 89.7%; Pred. No. 8e-170;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY      1 MAPSGLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKRLA 60
```

Db 1 MPPSGRLRLPLLLPLLLWLTGPPAAGLSTCKTIDMELVKRRIRIAIRGOILSKRLA 60  
Qy 61 SPSSQGDVPPGPPLEAVLALYNSTRDRVAGESVEPEPEADYAKVETRLVMVESGNOI 120  
Db 61 SPSSQGEVPPGPPLEAVLALYNSTRDRVAGESABPEPEADYAKVETRLVMVETHNEI 120  
Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Db 121 YDKFKQSTHSIYMFNTSELREAVPEVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Qy 181 YLSNRLAPSDEPWLSTFWTGVVROWLTRREALEGFRLSAHSSSDSKONTLHVEINGFN 240  
Db 181 YLSNRLAPSDEPWLSTFWTGVVROWLSRGGEIEGFRLSAHSCSDSDNTLQVDINGFT 240  
Qy 241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDPAASLALDTN 300  
Db 241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRRHR-----ALDTN 283  
Qy 301 YCFSTENKCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
Db 284 YCFSTENKCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
Qy 361 NOHNPASAAAPCCVQALPEPLPIVYVGRKPKVQQLSNMIVRSCKCS 407  
Db 344 NOHNPASAAAPCCVQALPEPLPIVYVGRKPKVQQLSNMIVRSCKCS 390

RESULT 5  
US-10-276-947-1  
; Sequence 1, Application US/10276947  
; Publication No. US20030176315A1  
; GENERAL INFORMATION:  
; APPLICANT: GLAXO GROUP LIMITED  
; TITLE OF INVENTION: MEDICAL PRODUCTS  
; FILE REFERENCE: PG3949  
; CURRENT APPLICATION NUMBER: US/10/276.947  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: GB 0012991.6  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: GB 0100286.4  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-276-947-1

Query Match 88.8%; Score 1916.5; DB 12; Length 390;  
Best Local Similarity 89.7%; Pred. No. 1.5e-169;  
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;  
Qy 1 MAPSGRLRLPLLLPLLLWLTGPPAAGLSTCKTIDMELVKRRIRIAIRGOILSKRLA 60  
Db 1 MPPSGRLRLPLLLPLLLWLTGPPAAGLSTCKTIDMELVKRRIRIAIRGOILSKRLA 60  
Qy 61 SPSSQGDVPPGPPLEAVLALYNSTRDRVAGESVEPEPEADYAKVETRLVMVESGNOI 120  
Db 61 SPSSQGEVPPGPPLEAVLALYNSTRDRVAGESABPEPEADYAKVETRLVMVETHNEI 120  
Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Db 121 YDKFKQSTHSIYMFNTSELREAVPEVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Qy 181 YLSNRLAPSDEPWLSTFWTGVVROWLTRREALEGFRLSAHSSSDSKONTLHVEINGFN 240  
Db 181 YLSNRLAPSDEPWLSTFWTGVVROWLSRGGEIEGFRLSAHSCSDSDNTLQVDINGFT 240  
Qy 241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDPAASLALDTN 300  
Db 241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRRHR-----ALDTN 283

Qy 301 YCFSTENKCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
Db 284 YCFSTENKCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
Qy 361 NOHNPASAAAPCCVQALPEPLPIVYVGRKPKVQQLSNMIVRSCKCS 407  
Db 344 NOHNPASAAAPCCVQALPEPLPIVYVGRKPKVQQLSNMIVRSCKCS 390

RESULT 6  
US-09-214-592-33  
; Sequence 33, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki, Cmotoo  
; APPLICANT: Shibata, Ckenji  
; APPLICANT: Sato, Cyasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214.592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 33  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: simian  
US-09-214-592-33

Query Match 88.7%; Score 1914.5; DB 11; Length 390;  
Best Local Similarity 89.4%; Pred. No. 2.3e-169;  
Matches 364; Conservative 10; Mismatches 16; Indels 17; Gaps 1;  
Qy 1 MAPSGRLRLPLLLPLLLWLTGPPAAGLSTCKTIDMELVKRRIRIAIRGOILSKRLA 60  
Db 1 MPPSGRLRLPLLLPLLLWLTGPPAAGLSTCKTIDMELVKRRIRIAIRGOILSKRLA 60  
Qy 61 SPSSQGDVPPGPPLEAVLALYNSTRDRVAGESVEPEPEADYAKVETRLVMVESGNOI 120  
Db 61 SPSSQGEVPPGPPLEAVLALYNSTRDRVAGESABPEPEADYAKVETRLVMVETHNEI 120  
Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Db 121 YDKFKQSTHSIYMFNTSELREAVPEVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Qy 181 YLSNRLAPSDEPWLSTFWTGVVROWLTRREALEGFRLSAHSSSDSKONTLHVEINGFN 240  
Db 181 YLSNRLAPSDEPWLSTFWTGVVROWLSRGGEIEGFRLSAHSCSDSDNTLQVDINGFT 240  
Qy 241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRRALDNTNSYPYDPAASLALDTN 300  
Db 241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRRHR-----ALDTN 283  
Qy 301 YCFSTENKCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
Db 284 YCFSTENKCCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
Qy 361 NOHNPASAAAPCCVQALPEPLPIVYVGRKPKVQQLSNMIVRSCKCS 407  
Db 344 NOHNPASAAAPCCVQALPEPLPIVYVGRKPKVQQLSNMIVRSCKCS 390

RESULT 7  
US-09-214-592-17  
; Sequence 17, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki, Cmotoo  
; APPLICANT: Shibata, Ckenji  
; APPLICANT: Sato, Cyasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD



QY 361 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 |||||  
 Db 344 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390  
 |||||

## RESULT 10

US-09-214-592-20

; Sequence 20, Application US/09214592A  
 ; Publication No. US20030027218A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yamasaki, Chotoo  
 ; APPLICANT: Shibata, Ckenji  
 ; APPLICANT: Sato, Ciyasufumi  
 ; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD  
 ; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
 ; FILE REFERENCE: 11060  
 ; CURRENT APPLICATION NUMBER: US/09/214, 592A  
 ; CURRENT FILING DATE: 1999-01-18  
 ; NUMBER OF SEQ ID NOS: 34  
 ; SOFTWARE:  
 ; SEQ ID NO 20  
 ; LENGTH: 390  
 ; TYPE: PRT  
 ; ORGANISM: murine  
 US-09-214-592-20

Query Match 85.0%; Score 1835.5; DB 11; Length 390;

Best Local Similarity 85.0%; Pred. No. 5.1e-162;

Matches 346; Conservative 15; Mismatches 29; Indels 17; Gaps 1;

QY 1 MAPSGRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRRIEARGOILSKRLA 60  
 |||||  
 Db 1 MPSPGRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRRIEARGOILSKRLA 60  
 |||||  
 QY 61 SPSPQGVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKETRVLMVSGNQI 120  
 |||||  
 Db 61 SPSPQGVPPGPLEAVLALYNSTRDRVAGESADPEPEADYAKETRVLMVDRNNAI 120  
 |||||  
 QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLKLKVEQHVLYKYNSDSWR 180  
 |||||  
 Db 121 YETKDISHSIYMFNTSDIREAVPEPVLLSRAELRLKLKVEQHVLYKYNSNSWR 180  
 |||||  
 QY 181 YLSNRLAPSDSPWLSFQVTVGVROWLRTREAIIEGFRLSAHSSSDSKDNTLHVEINGFN 240  
 |||||  
 Db 181 YLGNRLLTPTDTPWLSFQVTVGVROWLQOGDIOGFRSAHSCSDSKDNKLHVEINGIS 240  
 |||||  
 QY 241 SGRGDLATHGNRPFLLMATPLERAQHLSSRRRLDNTNSYPYDVPDYASLADTN 300  
 |||||  
 Db 241 PKRRGDLGTHDMNRPFLLMATPLERAQHLSSRRHR-----ALDTN 283  
 |||||  
 QY 301 YCFSTKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
 |||||  
 Db 284 YCFSTKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
 |||||  
 QY 361 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 |||||  
 Db 344 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390  
 |||||

## RESULT 11

US-09-214-592-23

; Sequence 23, Application US/09214592A  
 ; Publication No. US20030027218A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yamasaki, Chotoo  
 ; APPLICANT: Shibata, Ckenji  
 ; APPLICANT: Sato, Ciyasufumi  
 ; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD  
 ; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
 ; FILE REFERENCE: 11060  
 ; CURRENT APPLICATION NUMBER: US/09/214, 592A  
 ; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 23

; LENGTH: 390

; TYPE: PRT

; ORGANISM: rat

US-09-214-592-23

Query Match 85.0%; Score 1835.5; DB 11; Length 390;

Best Local Similarity 85.0%; Pred. No. 5.1e-162;

Matches 346; Conservative 14; Mismatches 30; Indels 17; Gaps 1;

QY 1 MAPSGRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRRIEARGOILSKRLA 60  
 |||||  
 Db 1 MPSPGRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKKRRIEARGOILSKRLA 60  
 |||||  
 QY 61 SPSPQGVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKETRVLMVSGNQI 120  
 |||||  
 Db 61 SPSPQGVPPGPLEAVLALYNSTRDRVAGESADPEPEADYAKETRVLMVDRNNAI 120  
 |||||  
 QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLKLKVEQHVLYKYNSDSWR 180  
 |||||  
 Db 121 YDKTKDITHSIYMFNTSDIREAVPEPVLLSRAELRLKLKVEQHVLYKYNSNSWR 180  
 |||||  
 QY 181 YLSNRLAPSDSPWLSFQVTVGVROWLRTREAIIEGFRLSAHSSSDSKDNTLHVEINGFN 240  
 |||||  
 Db 181 YLGNRLLTPTDTPWLSFQVTVGVROWLQOGDIOGFRSAHSCSDSKDNKLHVEINGIS 240  
 |||||  
 QY 241 SGRGDLATHGNRPFLLMATPLERAQHLSSRRRLDNTNSYPYDVPDYASLADTN 300  
 |||||  
 Db 241 PKRRGDLGTHDMNRPFLLMATPLERAQHLSSRRHR-----ALDTN 283  
 |||||  
 QY 301 YCFSTKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
 |||||  
 Db 284 YCFSTKNCVRLQYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
 |||||  
 QY 361 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 |||||  
 Db 344 NOHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390  
 |||||

## RESULT 12

US-09-214-592-25

; Sequence 25, Application US/09214592A  
 ; Publication No. US20030027218A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yamasaki, Chotoo  
 ; APPLICANT: Shibata, Ckenji  
 ; APPLICANT: Sato, Ciyasufumi  
 ; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD  
 ; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
 ; FILE REFERENCE: 11060  
 ; CURRENT APPLICATION NUMBER: US/09/214, 592A  
 ; CURRENT FILING DATE: 1999-01-18  
 ; NUMBER OF SEQ ID NOS: 34  
 ; SOFTWARE:  
 ; SEQ ID NO 25  
 ; LENGTH: 315  
 ; TYPE: PRT  
 ; ORGANISM: bovine  
 US-09-214-592-25

Query Match 72.8%; Score 1572.5; DB 11; Length 315;

Best Local Similarity 88.9%; Pred. No. 1e-137;

Matches 295; Conservative 9; Mismatches 11; Indels 17; Gaps 1;

QY 76 AVLALYNSTRDRVAGESVEPEPEADYAKETRVLMVSGNQIYDKFKGTPHSLYALF 135  
 |||||  
 Db 1 AILALYNSTRDRVAGESAETPEPEADYAKETRVLMVEYGNKYDKMKSSSHSIYMF 60  
 |||||  
 QY 136 NTSELREAVPEPVLLSRAELRLKLKVEQHVLYKYNSDSWRYSNRLAPSUSPEW 195  
 |||||  
 Db 61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVLYKYNSNSWRYSNRLAPSDSPW 120  
 |||||

QY 196 LSFDTGVVROWLTRREAEIEGFRLSAHSDDSKDNTLHVEINGFNSGRGDLATIHGMNR 255  
DB 121 LSFDTGVVROWLTRREAEIEGFRLSAHSDDSKDNTLQVINGFSSGRGDLATIHGMNR 180  
QY 256 PFLMATPLERAQHLHSSRRALDTNSYPVDPDYASLALDTNYCFSSSTEKNCCVROL 315  
DB 181 PFLMATPLERAQHLHSSRRR-----ALDTNYCFSSSTEKNCCVROL 223  
QY 316 YIDFRKDLGKWNTHPKGYHNPCLGCPYIWSLDTQYSKVLALYNQHNPGASAPCCVP 375  
DB 224 YIDFRKDLGKWNTHPKGYHNPCLGCPYIWSLDTQYSKVLALYNQHNPGASAPCCVP 283  
QY 376 QALEPLPIVYVGRKPKVQSLNMIVRSCKS 407  
DB 284 QALEPLPIVYVGRKPKVQSLNMIVRSCKS 315  
RESULT 13  
US-09-756-283A-20  
; Sequence 20, Application US/09756283A  
; Patent No. US20020151478A1  
; GENERAL INFORMATION:  
; APPLICANT: Chernajovsky, Yuti  
; APPLICANT: Dreja, Hanna Stina  
; APPLICANT: Adams, Gillian  
; TITLE OF INVENTION: Latent Fusion Protein  
; FILE REFERENCE: 0623.1000000  
; CURRENT APPLICATION NUMBER: US/09/756,283A  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: LAP-mifn construct  
US-09-756-283A-20

Query Match 58.5%; Score 1262.5; DB 10; Length 455;  
Best Local Similarity 90.5%; Pred. No. 1.1e-108;  
Matches 248; Conservative 10; Mismatches 15; Indels 1; Gaps 1;  
QY 1 MAPSGLRLPLPLPLLWLVITPGPAAGLSTCKTIDMELVKRKRIEARGOILSKRLA 60  
DB 1 MPESGLRLPLPLPLLWLVITPGPAAGLSTCKTIDMELVKRKRIEARGOILSKRLA 60  
QY 61 SPSPQGDVPPGLPEAVLALYNSTRORVAGESVEPEPEADYAKVETRVLMVESGNOI 120  
DB 61 SPSPQGEVPPGLPEAVLALYNSTRORVAGESAEPEPEADYAKVETRVLMVETHNEI 120  
QY 121 YDKFKTGHSLYMLFNTSELREAVPEPVLLSRAELRL-RLKLVQHOHVELYOKYSNDSW 179  
DB 121 YDKFKTGHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQHOHVELYOKYSNDSW 180  
QY 180 RYLSNRLAPSDPEWLSFDVTGVVROWLTRRDAIEGFRLSAHSDDSKDNTLHVEINGF 239  
DB 181 RYLSNRLAPSDPEWLSFDVTGVVROWLTRRDAIEGFRLSAHSDDSKDNTLQVDINGF 240  
QY 240 NSGRGDLATIHGMNRPFLMATPLERAQHLHS 273  
DB 241 TTGRRGDLATIHGMNRPFLMATPLERAQHLHS 274

RESULT 14  
US-09-756-283A-22  
; Sequence 22, Application US/09756283A  
; Patent No. US20020151478A1  
; GENERAL INFORMATION:  
; APPLICANT: Chernajovsky, Yuti  
; APPLICANT: Dreja, Hanna Stina  
; APPLICANT: Adams, Gillian

; TITLE OF INVENTION: Latent Fusion Protein  
; FILE REFERENCE: 0623.1000000  
; CURRENT APPLICATION NUMBER: US/09/756,283A  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: mifn -LAP construct  
US-09-756-283A-22  
Query Match 52.8%; Score 1139; DB 10; Length 447;  
Best Local Similarity 87.0%; Pred. No. 3.4e-97;  
Matches 228; Conservative 10; Mismatches 18; Indels 6; Gaps 3;  
QY 14 PL-LWLLVLTGPAAGLSTCKTIDMELVKRKRIEARGOILSKRLASPPSQGDVPPGP 72  
DB 190 PLGLW----AGGSAALSTCKTIDMELVKRKRIEARGOILSKRLASPPSQGEVPPGP 245  
QY 73 LPEAVLALYNSTRORVAGESVEPEPEADYAKVETRVLMVESGNOIYDKFKTGHSLY 132  
DB 246 LPEAVLALYNSTRORVAGESAEPEPEADYAKVETRVLMVETHNEIYDKFKQSTHSIY 305  
QY 133 MLFNTSELREAVPEPVLLSRAELRL-RLKLVQHOHVELYOKYSNDSWRYLSNRLAPSD 191  
DB 306 MFFNTSELREAVPEPVLLSRAELRLRLKLVQHOHVELYOKYSNNRMRYLSNRLAPSD 365  
QY 192 SPWLSFDVTGVVROWLTRRDAIEGFRLSAHSDDSKDNTLHVEINGFNSGRGDLATIH 251  
DB 366 SPWLSFDVTGVVROWLTRRDAIEGFRLSAHSDDSKDNTLQVDINGFTTGRGDLATIH 425  
QY 252 GMNRPFLMATPLERAQHLHS 273  
DB 426 GMNRPFLMATPLERAQHLHS 447

RESULT 15  
US-09-214-592-32  
; Sequence 32, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki, Chotoo  
; APPLICANT: Shibata, Kenji  
; APPLICANT: Sato, Chisufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 32  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: chicken  
US-09-214-592-32

Query Match 48.5%; Score 1047; DB 11; Length 373;  
Best Local Similarity 52.9%; Pred. No. 9.2e-89;  
Matches 207; Conservative 52; Mismatches 100; Indels 32; Gaps 8;  
QY 30 LSTCKTIDMELVKRKRIEARGOILSKRLASPPSQGDVPPGLPEAVLALYNSTRDVA 89  
DB 2 LSTCKTIDMELVKRKRIEARGOILSKRLASPPSQGDVPPGLPEAVLALYNSTRDVA 61  
QY 90 GES-VEPEPEADYAKVETRVLMVESGNOIYDKFKTGHSLYMLFNTSELREAVPEPV 148  
DB 62 QBARLRPPDGDDEYKAKELRIPMETTWGAMENHQPSHSIFVFVNSRRRG-GRPT 120  
QY 149 LLSRAELRLRLKLVQHOHVELYOKYSNDSWRYLSNRLAPSDPEWLSFDVTGV 203



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: October 7, 2003, 17:59:17 ; Search time 42 Seconds  
(without alignments)  
931.920 Million cell updates/sec

Title: US-10-017-372E-37  
Perfect score: 2159  
Sequence: 1 MAPSGRLRLPLLLPLMLLV.....GRKPKVEQLSNMIVRSKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2033.5	94.2	390	2 A27512	transforming growt
2	2002	92.7	391	2 S01413	transforming growt
3	1929.5	89.4	390	2 I46463	transforming growt
4	1916.5	88.8	390	1 WPHU2	transforming growt
5	1914.5	88.7	390	2 A26960	transforming growt
6	1893.5	87.7	390	2 JC4023	transforming growt
7	1835.5	85.0	390	1 WPM52	transforming growt
8	1835.5	85.0	390	2 S10219	transforming growt
9	1572.5	72.8	315	2 A40057	transforming growt
10	1047	48.5	373	2 A41918	transforming growt
11	871.5	40.4	412	2 A34939	transforming growt
12	868.5	40.2	382	2 B61036	transforming growt
13	834	38.6	409	2 S01825	transforming growt
14	827.5	38.3	410	2 A41397	transforming growt
15	826.5	38.3	410	2 A55706	transforming growt
16	824.5	38.2	412	2 A36189	transforming growt
17	824	38.2	414	1 WPM52	transforming growt
18	809	37.5	414	1 WPM52	transforming growt
19	809	37.5	414	2 A31249	transforming growt
20	808.5	37.4	413	1 WFXLB2	transforming growt
21	808	37.4	412	2 A39489	transforming growt
22	793	36.7	442	2 B31249	transforming growt
23	691.5	32.0	130	2 I48196	transforming growt
24	482	22.3	112	2 A61439	transforming growt
25	275	12.7	425	2 I47072	inhibin beta-A cha
26	272.5	12.6	424	1 WPGGA	inhibin beta-A cha
27	270.5	12.5	424	1 S31440	inhibin beta-A cha
28	270	12.5	425	1 S50898	inhibin beta-A cha
29	268.5	12.4	426	1 B24248	inhibin beta-A cha

30	266.5	12.3	424	1 B40905	inhibin beta-A cha
31	256.5	11.9	513	1 BMHU6	bone morphogenetic
32	254.5	11.8	398	2 JH0688	bone morphogenetic
33	253.5	11.7	413	2 JC4862	activin beta-A cha
34	253	11.7	394	2 S45355	bone morphogenetic
35	253	11.7	398	2 JH0687	bone morphogenetic
36	252	11.7	396	1 BMHU2	bone morphogenetic
37	246	11.4	510	2 A54798	Vg-1-related prote
38	244.5	11.3	393	2 S37073	bone morphogenetic
39	243	11.3	455	2 A43918	TGF-beta-related p
40	238	11.0	350	2 JC5241	activin beta E cha
41	234.5	10.9	402	2 A45056	osteoegenic protein
42	231	10.7	420	2 I49541	bone morphogenetic
43	228	10.6	454	1 BMHU5	bone morphogenetic
44	227	10.5	367	2 JC4151	activin beta D cha
45	225.5	10.4	430	2 JQ1184	osteoegenic protein

## ALIGNMENTS

RESULT 1  
A27512 transforming growth factor beta-1 precursor - pig  
N:Alternate names: TGF-beta  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1988 #sequence.revision 05-Jun-1988 #text\_change 16-Jul-1999  
C:Accession: A27512; A26356; I46657  
R:Derynck, R.; Rhee, L.  
Nucleic Acids Res. 15, 1387, 1987  
A:Title: Sequence of the porcine transforming growth factor-beta precursor.  
A:Reference number: A27512; MUID:87174844; PMID:3470708  
A:Accession: A27512  
A:Molecule type: mRNA  
A:Residues: 1-390 <DER>  
R:Chelifetz, S.; Weatherbee, J.A.; Tsang, M.L.S.; Anderson, J.K.; Mole, J.E.; Lucas, R.  
Cell 48, 409-415, 1987  
A:Title: The transforming growth factor-beta system, a complex pattern of cross-react  
A:Reference number: A90890; MUID:87102890; PMID:2879635  
A:Accession: A26356  
A:Molecule type: protein  
A:Residues: 279-322 <CHE>  
R:Kondalah, P.; Van Oberghe-Schilling, E.; Ludwig, R.L.; Dhar, R.; Sporn, M.B.; Rob  
J. Biol. Chem. 263, 18313-18317, 1988  
A:Title: cDNA cloning of porcine transforming growth factor-beta 1 mRNAs. Evidence fo  
A:Reference number: I46657; MUID:89054010; PMID:2461367  
A:Accession: I46657  
A>Status: preliminary: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-390 <KON>  
A:Cross-references: GB:M23703; NID:g755044; PIDN:AAA64616.1; PID:g755045  
C:Genetics:  
A:Gene: TGF-beta-1  
C:Superfamily: inhibin  
C:Keywords: growth factor

Query Match		94.2%	Score 2033.5;	DB 2;	Length 390;
Best Local Similarity		95.1%	Pred. No. 6.9e-156;		
Matches 387;		Conservative	0;	Mismatches	3;
				Indels	17;
				Gaps	1;
Qy	1	MAPSGRLRLPLLLPLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIEARGQILSKRLA	60		
Db	1	MAPSGRLRLPLLLPLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIEARGQILSKRLA	60		
Qy	61	SPSQGVDPGCPPLPEAVLALYNSTRVACSEVPEPEPEADYAKEVTRVLAVESNQI	120		
Db	61	SPSQGVDPGCPPLPEAVLALYNSTRVACSEVPEPEPEADYAKEVTRVLAVESNQI	120		
Qy	121	YDFKGTGPHSLYMLFNTSELREAVPEPVLLSRRAELRLRLKLVQEHVELYQKYSNDSWR	180		
Db	121	YDFKGTGPHSLYMLFNTSELREAVPEPVLLSRRAELRLRLKLVQEHVELYQKYSNDSWR	180		
Qy	181	YLSNRLAPSDSPSEWLSFDVTGVVROWLTRREATIEGFRLSAHSSSDSKDNTHLVEINGFN	240		



Db 181 YLSNRLAPSDSPWLSFDVTGVVQWLTREAEGLSAHSCSDKNTLHVEINGFN 240  
 Qy 241 SGRGDLATIHGMNRPFLLLNATPLERAQHLHSHRRALDNTSYDPDYASLALDTN 300  
 Db 241 SGRGDLATIHGMNRPFLLLNATPLERAQHLHSHRRR-----ALDNTN 283  
 Qy 301 YCFSTSEKNCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
 Db 284 YCFSTSEKNCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
 Qy 361 QNHNPASAPCCVPQALEPPIVYVYGRPKVQOLSNMIVRSCKCS 407  
 Db 344 QNHNPASAPCCVPQALEPPIVYVYGRPKVQOLSNMIVRSCKCS 390  
 RESULT 2  
 S01413  
 transforming growth factor beta-1 precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 24-Nov-1999  
 C:Accession: S01413  
 R:Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.  
 Nucleic Acids Res. 16, 8730, 1988  
 A:Title: Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1).  
 A:Reference number: S01413; MUID:83335639; PMID:3166520  
 A:Accession: S01413  
 A:Molecule type: DNA  
 A:Residues: 1-391 <JAK>  
 A:Cross-references: EMBL:X12373; NID:963808; PIDN:CAA30933.1; PID:g63809  
 C:Superfamily: inhibin  
 C:Keywords: growth factor

Query Match 92.7%; Score 2002; DB 2; Length 391;  
 Best Local Similarity 94.1%; Pred. No. 2.4e-153;  
 Matches 384; Conservative 0; Mismatches 6; Indels 18; Gaps 2;  
 Qy 1 MAPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGOILSKRLA 60  
 Db 1 MPPSGGLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGOILSKRLA 60  
 Qy 61 SPPSGQGVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKEVTRVLMVESQNI 120  
 Db 61 SPPSGQGVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKEVTRVLMVESQNI 120  
 Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVKEQHVLYQKYSNDSWR 180  
 Db 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVKEQHVLYQKYSNDSWG 180  
 Qy 181 YLSNRLAPSDSPWLSFDVTGVVQWLTREAEGLSAHSSSDSKDNTLHVEIN-GF 239  
 Db 181 YLSNRLAPSDSPWLSFDVTGVVQWLTREAEGLSAHSCSDSKDNTLHVEINAGF 240  
 Qy 240 NSGRGDLATIHGMNRPFLLLNATPLERAQHLHSHRRALDNTSYDPDYASLALDT 299  
 Db 241 NSGRGDLATIHGMNRPFLLLNATPLERAQHLHSHRRR-----ALDT 283  
 Qy 300 NYCFSTSEKNCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLAL 359  
 Db 284 NYCFSTSEKNCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLAL 343  
 Qy 360 QNHNPASAPCCVPQALEPPIVYVYGRPKVQOLSNMIVRSCKCS 407  
 Db 344 QNHNPASAPCCVPQALEPPIVYVYGRPKVQOLSNMIVRSCKCS 391

RESULT 3  
 I46463  
 transforming growth factor beta-1 - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 24-Nov-1999  
 C:Accession: I46463; S45115  
 R:Woodall, C.J.; McLaren, L.J.; Watt, N.J.

Gene 150, 371-373, 1994  
 A:Title: Sequence and chromosomal localisation of the gene encoding ovine latent tran  
 A:Reference number: I46463; MUID:95121932; PMID:7821809  
 A:Accession: I46463  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-390 <WOO>  
 A:Cross-references: EMBL:X76916; NID:g496648; PIDN:CAA54242.1; PID:g496649  
 A:Note: submitted to the EMBL Data Library, December 1993  
 C:Superfamily: inhibin

Query Match 89.4%; Score 1929.5; DB 2; Length 390;  
 Best Local Similarity 89.7%; Pred. No. 1.6e-147;  
 Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;  
 Qy 1 MAPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGOILSKRLA 60  
 Db 1 MPPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGOILSKRLA 60  
 Qy 61 SPPSGQGVPPGPLEAVLALYNSTRDRVAGESVEPEPEADYAKEVTRVLMVESQNI 120  
 Db 61 SPPSGQGVPPGPLEAVLALYNSTRDRVAGESAEPEPEADYAKEVTRVLMVEYGNKI 120  
 Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRALRLRLKLVKEQHVLYQKYSNDSWR 180  
 Db 121 YDKMKSSSHSIYFFNTSELREAVPEPVLLSRADVRLRLKLVKEQHVLYQKYSNNSWR 180  
 Qy 181 YLSNRLAPSDSPWLSFDVTGVVQWLTREAEGLSAHSSSDSKDNTLHVEINGFN 240  
 Db 181 YLSNRLAPSDSPWLSFDVTGVVQWLTREAEGLSAHSCSDSKDNTLQVDINGFS 240  
 Qy 241 SGRGDLATIHGMNRPFLLLMATPLERAQHLHSHRRALDNTSYDPDYASLALDTN 300  
 Db 241 SGRGDLATIHGMNRPFLLLMATPLERAQHLHSHRRR-----ALDIN 283  
 Qy 301 YCFSTSEKNCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360  
 Db 284 YCFSTSEKNCVROLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
 Qy 361 QNHNPASAPCCVPQALEPPIVYVYGRPKVQOLSNMIVRSCKCS 407  
 Db 344 QNHNPASAPCCVPQALEPPIVYVYGRPKVQOLSNMIVRSCKCS 390

RESULT 4  
 WPHU2  
 transforming growth factor beta-1 precursor [validated]; - human  
 N:Alternate names: growth-inhibitory factor; TGF type 2; TGF-beta  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Feb-1986 #sequence\_revision 19-Oct-1995 #text\_change 08-Dec-2000  
 C:Accession: A27513; A01395; A22290; I59664; S53444  
 R:Derynck, R.; Rhee, L.; Chen, E.Y.; Van Tilburg, A.  
 Nucleic Acids Res. 15, 3188-3189, 1987  
 A:Title: Intronic structure of the human transforming growth factor-beta precursor  
 A:Reference number: A27513; MUID:87174845; PMID:3470709  
 A:Accession: A27513  
 A:Molecule type: DNA  
 A:Residues: 1-390 <DER>  
 A:Cross-references: GB:X05839; GB:Y00112; NID:g37097; PIDN:CAA29283.1; PID:g1212989  
 R:Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Eaton, D.H.; Bell, J.R.; Assoian, R.K.; Rob  
 Nature 316, 701-705, 1985  
 A:Title: Human transforming growth factor-beta complementary DNA sequence and express  
 A:Reference number: A01395; MUID:85296301; PMID:3861940  
 A:Accession: A01395  
 A:Molecule type: mRNA  
 A:Residues: 1-9,'p',11-24,'p',26-159,'R',160-390 <DE2>  
 A:Cross-references: GB:X02812; GB:J05114; NID:g37092; PIDN:CAA26580.1; PID:g37093  
 A:Note: the authors suggest that residues 8-23 could represent the hydrophobic core o  
 R:Massague, J.; Like, B.  
 J. Biol. Chem. 260, 2636-2645, 1985  
 A:Title: Cellular receptors for type beta transforming growth factor. Ligand binding  
 A:Reference number: A22290; MUID:85131019; PMID:2982829  
 A:Accession: A22290

A:Molecule type: protein  
A:Residues: 279-395 'XX', 298-301 <MAS>  
R:Urushizaki, Y.; Nitsui, Y.; Terui, T.; Koshida, Y.; Mahara, K.; Kohgo, Y.; Urushizaki, Tumor Res. 22, 41-55, 1987  
A:Title: Cloning and expression of the gene for human transforming growth factor-beta in A:Reference number: I59664  
A:Accession: I59664  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 279-390 <RES>  
A:Cross-references: GB:M38449; NID:g339557; PIDN:AAA36735.1; PID:g339558  
R:Stam, K.; Stewart, A.A.; Qu, G.Y.; Iwata, K.K.; Fenyo, D.; Chait, B.T.; Marshak, D.R. Biochem. J. 305, 87-92, 1995  
A:Title: Physical and biological characterization of a growth-inhibitory activity purified A:Reference number: S53444; MUID:95126934; PMID:7826358  
A:Accession: S53444  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 279-297 <STA>  
C:Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide C:Genetics:  
A:Gene: GDB:TGFBI; TGFBI  
A:Cross-references: GDB:120729; OMIM:190180  
A:Map position: 19q13.2-19q13.2  
C:Superfamily: inhibin  
C:Keywords: glycoprotein; growth factor; homodimer; mitogen; transformation  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-278/Domain: propeptide #status predicted <PRO>  
F:244-246/Region: cell attachment (R-G-D) motif  
F:279-390/Product: transforming growth factor beta-1 #status experimental <MAT>  
F:82.136.176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.8%; Score 1916.5; DB 1; Length 390;  
Best Local Similarity 89.7%; Pred. No. 1.8e-146;  
Matches 365; Conservative 10; Mismatches 17; Indels 17; Gaps 1;

Qy 1 MAPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
Db 1 MPSPGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
Qy 61 SPSPGQDVPPLPPLPAVLALYNSTRDRVAGESVEPEPEADYAKETRVLMVESGNOI 120  
Db 61 SPSPGQEVPPGPIPAVLALYNSTRDRVAGESAEPEPEADYAKETRVLMVETHNEI 120  
Qy 121 YDKFKGTPLSHLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Db 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNNSWR 180  
Qy 181 YLSNRLAPSDSPWLSFDVTGVVQWLTRRAIEGFRLSAHSSDSKNTLHVEINGFN 240  
Db 181 YLSNRLAPSDSPWLSFDVTGVVQWLTRRAIEGFRLSAHSSDSKNTLQVDINGFT 240  
Qy 241 SGRRGDLATIHGMNRPFLMATPLERAQHLHSSRRRALDTSYDVPDYASLADTN 300  
Db 241 TGRGDLATIHGMNRPFLMATPLERAQHLHSSRRR-----ALDIN 283  
Qy 301 YCFSTENKCCVROLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLY 360  
Db 284 YCFSTENKCCVROLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLY 343  
Qy 361 NOHNPASAPCCVPOALEPLIVYVGRKPKVEQLSNMIVRSCKS 407  
Db 344 NOHNPASAPCCVPOALEPLIVYVGRKPKVEQLSNMIVRSCKS 390

RESULT 6  
JC4023  
transforming growth factor beta-1 - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 24-Nov-1999  
C:Accession: JC4023  
R:Manning, A.M.; Auchampach, J.A.; Drong, R.F.; Slightom, J.L. Gene 155, 307-308, 1995  
A:Title: Cloning of a canine cDNA homologous to the human transforming growth factor- A:Reference number: JC4023; MUID:95237630; PMID:772110  
A:Accession: JC4023  
A:Molecule type: mRNA  
A:Residues: 1-390 <MAN>  
A:Cross-references: GB:L34956; NID:g516071; PIDN:AAA51458.1; PID:g516072  
C:Comment: This factor plays a multifunctional role as a regulator of mammalian cell C:Genetics:  
A:Gene: tqf-beta1  
C:Superfamily: inhibin  
C:Keywords: growth factor; transforming protein  
F:288-390/Product: transforming growth factor beta 1 #status predicted <MAT>

Query Match 87.7%; Score 1893.5; DB 2; Length 390;  
Best Local Similarity 88.7%; Pred. No. 1.3e-144;  
Matches 361; Conservative 12; Mismatches 17; Indels 17; Gaps 1;

Qy 1 MAPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
Db 1 MPSPGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLS 60

A:Title: Cloning and sequence analysis of simian transforming growth factor-beta cDNA A:Reference number: A26960; MUID:87246074; PMID:3474130  
A:Accession: A26960  
A:Molecule type: mRNA  
A:Residues: 1-390 <SHA>  
A:Cross-references: GB:M16658; NID:g176552; PIDN:AAA35369.1; PID:g176553  
C:Superfamily: inhibin  
C:Keywords: growth factor  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-390/Product: transforming growth factor beta #status predicted <MAT>

Query Match 88.7%; Score 1914.5; DB 2; Length 390;  
Best Local Similarity 89.4%; Pred. No. 2.6e-146;  
Matches 364; Conservative 10; Mismatches 16; Indels 17; Gaps 1;

Qy 1 MAPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
Db 1 MPSPGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
Qy 61 SPSPGQDVPPLPPLPAVLALYNSTRDRVAGESVEPEPEADYAKETRVLMVESGNOI 120  
Db 61 SPSPGQEVPPGPIPAVLALYNSTRDRVAGESAEPEPEADYAKETRVLMVETHNEI 120  
Qy 121 YDKFKGTPLSHLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNDSWR 180  
Db 121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVQEHVELYQKYSNNSWR 180  
Qy 181 YLSNRLAPSDSPWLSFDVTGVVQWLTRRAIEGFRLSAHSSDSKNTLHVEINGFN 240  
Db 181 YLSNRLAPSDSPWLSFDVTGVVQWLTRRAIEGFRLSAHSSDSKNTLQVDINGFT 240  
Qy 241 SGRRGDLATIHGMNRPFLMATPLERAQHLHSSRRRALDTSYDVPDYASLADTN 300  
Db 241 TGRGDLATIHGMNRPFLMATPLERAQHLHSSRRR-----ALDIN 283  
Qy 301 YCFSTENKCCVROLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLY 360  
Db 284 YCFSTENKCCVROLYIDFRKDLGKWIHEPKGYHANFCLGCPYIWSLDTQYSKVLY 343  
Qy 361 NOHNPASAPCCVPOALEPLIVYVGRKPKVEQLSNMIVRSCKS 407  
Db 344 NOHNPASAPCCVPOALEPLIVYVGRKPKVEQLSNMIVRSCKS 390

RESULT 6  
JC4023  
transforming growth factor beta-1 - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 24-Nov-1999  
C:Accession: JC4023  
R:Manning, A.M.; Auchampach, J.A.; Drong, R.F.; Slightom, J.L. Gene 155, 307-308, 1995  
A:Title: Cloning of a canine cDNA homologous to the human transforming growth factor- A:Reference number: JC4023; MUID:95237630; PMID:772110  
A:Accession: JC4023  
A:Molecule type: mRNA  
A:Residues: 1-390 <MAN>  
A:Cross-references: GB:L34956; NID:g516071; PIDN:AAA51458.1; PID:g516072  
C:Comment: This factor plays a multifunctional role as a regulator of mammalian cell C:Genetics:  
A:Gene: tqf-beta1  
C:Superfamily: inhibin  
C:Keywords: growth factor; transforming protein  
F:288-390/Product: transforming growth factor beta 1 #status predicted <MAT>

Query Match 87.7%; Score 1893.5; DB 2; Length 390;  
Best Local Similarity 88.7%; Pred. No. 1.3e-144;  
Matches 361; Conservative 12; Mismatches 17; Indels 17; Gaps 1;

Qy 1 MAPSGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLA 60  
Db 1 MPSPGRLRLPLLLPLLLVLTGPRPAAGLSTCKTIDMELVKRKRIEIRGQILSKRLS 60



Db 284 YCFSTSEKNCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343  
 QY 361 NQHNPGASAAAPCCVPQALPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 Db 344 NQHNPGASAPCCVPQALPLPIVYVGRKPKVEQLSNMIVRSCKCS 390  
 RESULT 9  
 A40057  
 transforming growth factor beta-1 precursor - bovine (fragment)  
 N:Alternate names: beta-TGF; cartilage-inducing factor-A; EGF-dependent TGF or dEGF; MGF  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 28-Feb-1992 \*sequence\_revision 28-Feb-1992 \*text\_change 16-Jul-1999  
 C:Accession: A40057; A42320; A05284; A24322; B61439  
 R:Van Obberghen-Schilling, E.; Kondaiah, P.; Ludwig, R.L.; Sporn, M.B.; Baker, C.C.  
 Mol. Endocrinol. 1, 693-698, 1987  
 A:Title: Complementary deoxyribonucleic acid cloning of bovine transforming growth factor  
 A:Reference number: A40057; MUID:91042552; PMID:3153459  
 A:Accession: A40057  
 A:Molecule type: mRNA  
 A:Residues: 1-315 <VAN>  
 A:CROSS-references: GB:M36271; NID:q163747; PIDN:AAA30778.1; PID:q163748  
 R:Ogawa, Y.; Schmidt, D.K.; Dasch, J.R.; Chang, R.J.; Glasor, C.B.  
 J. Biol. Chem. 267, 2325-2328, 1992  
 A:Title: Purification and characterization of transforming growth factor-beta2.3 and -beta  
 A:Reference number: A42320; MUID:92129307; PMID:1733936  
 A:Accession: A42320  
 A:Molecule type: protein  
 A:Residues: 204-209, 'X', 211-217 <OGA>  
 R:Roberts, A.B.; Anzano, M.A.; Meyers, C.A.; Wideman, J.; Blacher, R.; Pan, Y.C.E.; Stein  
 Blochmistry 22, 5692-5698, 1983  
 A:Title: Purification and properties of a type beta transforming growth factor from bovi  
 A:Reference number: A05284; MUID:34104793; PMID:6607069  
 A:Accession: A05284  
 A:Molecule type: protein  
 A:Residues: 204-218 <ROB>  
 R:Sevedin, S.M.; Thompson, A.X.; Bentz, H.; Rosen, D.M.; McPherson, J.M.; Conti, A.; Sie  
 J. Biol. Chem. 261, 5693-5695, 1936  
 A:Title: Cartilage-inducing factor-A. Apparent identity to transforming growth factor-be  
 A:Reference number: A24322; MUID:36195954; PMID:3754555  
 A:Accession: A24322  
 A:Molecule type: protein  
 A:Residues: 204-233 <SVY>  
 R:Jin, Y.; Cox, D.A.; Knecht, R.; Raschdorf, F.; Cerletti, N.  
 J. Protein Chem. 10, 565-575, 1991  
 A:Title: Separation, purification, and sequence identification of TGF-beta1 and TGF-beta  
 A:Reference number: A61439; MUID:92189724; PMID:1799413  
 A:Accession: B61439  
 A:Molecule type: protein  
 A:Residues: 204-209, 'X', 211-217, 'XX', 220-232 <JIN>  
 C:Comment: This polypeptide is composed of two polypeptide chains cross-linked by disulf  
 C:Comment: Type II TGF does not bind to the EGF receptor and lacks intrinsic mitogenic a  
 ion. Cells grown in monolayer do not respond in a similar manner to these growth factors  
 C:Superfamily: Inhibin  
 C:Keywords: glycoprotein; growth factor; heterodimer  
 F:204-315/Product: transforming growth factor beta-1 \*status experimental  
 F:7,61,101/Binding site: carbohydrate (Asn) (covalent) \*status predicted  
 Query Match 72.8%; Score 1572.5; DB 2; Length 315;  
 Best Local Similarity 88.9%; Pred. No. 6.6e-119;  
 Matches 295; Conservative 9; Mismatches 11; Indels 17; Gaps 1;  
 QY 76 AVIALYNSTRDRVAGESVEPEADYAKETVRLVMEGNOIYDKFKGTPHSLYMLF 135  
 Db 1 AIALYNSTRDRVAGESAETPEADYAKETVRLVMEYGNRIYDKMKSSSHSYMF 60  
 QY 136 NTSSELRAVPEPVLRSRAELRLRLKLVQHQVELYQKYSNDSWRYLSNRLIAPSDSPEN 195  
 Db 61 NTSSELRAVPEPVLRSRAVRLRLRLKLVQHQVELYQKYSNNSWRYLSNRLIAPSDSPEN 120  
 QY 196 LSFDTGVVRQWLTRRAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGLATIHMNR 255  
 Db 121 LSFDTGVVRQWLTRREIEGFRLSAHCSDSKDNTLOVDINGFSSGRRGLATIHMNR 180

QY 256 PFLLLMATPLERAQHLHSSRRHRLDNTNSYPDYVDPYASLALDNTYCFSSSTKNCVQRL 315  
 Db 181 PFLLLMATPLERAQHLHSSRRHRLDNTNSYPDYVDPYASLALDNTYCFSSSTKNCVQRL 223  
 QY 316 YIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALYQHNPNGASAAAPCCVP 375  
 Db 224 YIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALYQHNPNGASAAAPCCVP 283  
 QY 376 QALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407  
 Db 284 QALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 315  
 RESULT 10  
 A41918  
 transforming growth factor beta-4 precursor - chicken (fragment)  
 N:Alternate names: TGF-beta 4  
 C:Species: Gallus gallus (chicken)  
 C:Date: 31-Dec-1993 \*sequence\_revision 31-Dec-1993 \*text\_change 16-Jul-1999  
 C:Accession: A41918; A34941; S03110  
 R:Burt, D.W.; Jakowlew, S.H.  
 Mol. Endocrinol. 6, 989-992, 1992  
 A:Title: Correction: a new interpretation of a chicken transforming growth factor-bet  
 A:Reference number: A41918; MUID:92357039; PMID:1353860  
 A:Accession: A41918  
 A:Molecule type: mRNA  
 A:Residues: 1-373 <BUR>  
 A:CROSS-references: GB:M31160; GB:X08012; GB:S41706; NID:gl262437; PIDN:AA805637.1; P  
 A:Note: sequence extracted from NCBI backbone (NCBIN:110186, NCBIIP:110187)  
 R:Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.  
 Mol. Endocrinol. 2, 1186-1195, 1988  
 A:Title: Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid  
 A:Reference number: A34941; MUID:89112198; PMID:2464131  
 A:Accession: A34941  
 A:Molecule type: mRNA  
 A:Residues: 1MDPM5GKSGSPWPCPTAWMSGRRATSSSCSTSSRVRAEVGGRL',122-209,'D',211-37  
 A:CROSS-references: EMBL:X08012  
 A:Note: this sequence has been corrected in A41918  
 C:Superfamily: Inhibin  
 C:Keywords: glycoprotein; growth factor  
 F:1/Domain: signal sequence (fragment) \*status predicted <SIG>  
 F:223-225/Region: cell attachment (R-G-D) motif  
 F:260-373/Product: transforming growth factor beta-4 \*status predicted <MAT>  
 F:54,109,153/Binding site: carbohydrate (Asn) (covalent) \*status predicted  
 Query Match 48.5%; Score 1047; DB 2; Length 373;  
 Best Local Similarity 52.9%; Pred. No. 1.7e-76;  
 Matches 207; Conservative 52; Mismatches 100; Indels 32; Gaps 8;  
 QY 30 LSTCKTIDMELVKKRIEAIHQILSKRLASPPSGDVPVPGPLPEAVLALYNSTRDVA 89  
 Db 2 LSTCORLDEAAKKRIEAVRGQILSKRLTAPPASETPPRPLPDVDRALYNSTQELLK 61  
 QY 90 GSGS-VEPEPEADYAKETVRLVMEGNOIYDKFKGTPHSLYMLNFTSELREAVPEPV 148  
 Db 62 QARLRPPDGPDEYWAKELRLPMTTWGMEHWPQSHSIFVFNVSRRRG-GRPT 120  
 QY 149 LLSRAELRLRLKLVK-----VEQHVELYQKYSNDSWRYLSNRLIAPSDSPENLSEFDTGV 203  
 Db 121 LLRAELRLRLQKAAADSAGTEQRLLEYQYGYNASWRYLHGRSVRAIADWLSFDTDA 180  
 QY 204 VQWLTRRAIEGFRLSAHSSSD---SKDNTLHVEINGFNSGRRGLATIHMNR--PFL 258  
 Db 181 VQWLGSSELLGVKFLSHCPCEMGEGHAEEMRIETEGFEQ-QRGDMQSIKAKRRVPV 239  
 QY 259 LLMATPLERAQHLHSSRRHRLDNTNSYPDYVDPYASLALDNTYCF--SSTEKNCCVQRL 316  
 Db 240 LAMALPAERANELSHARRR-----DLDTYCGFGTDFKNCVQRL 282  
 QY 317 IDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALYQHNPNGASAAAPCCVP 376

Db 283 IDFRKDLQMKWIHEPKGYMANFCMPCPYIWSADTQYIKVLYLNQHNPGASAAAPCCVPQ 342

QY 377 ALPLPIVTVYVGRKPKVEQLSNMIVRSCKCS 407

Db 343 TLDPLPIVTVYVGRNVRVEQLSNMIVRSCKCS 373

RESULT 11

A34939

transforming growth factor beta-3 precursor - chicken

C:Species: Gallus gallus (chicken)

C>Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 16-Jul-1999

C:Accession: A34939; S36125; S36124; I51181

R:Jakowlew, S.B.; Dillard, P.J.; Kondaiah, P.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 2, 747-755, 1988

A:Title: Complementary deoxyribonucleic acid cloning of a novel transforming growth factor

A:Reference number: A34939; MUID:89096966; PMID:3211158

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-412 <JAK>

A:Cross-references: GB:M31154; NID:g212758; PIDN:AAA49089.1; PID:g212759

R:Burt, D.W.; Paton, I.R.; Dey, B.X.

J. Mol. Endocrinol. 7, 175-183, 1991

A:Title: Comparative analysis of human and chicken transforming growth factor-beta-2 and

A:Reference number: S25850; MUID:92134496; PMID:1840616

A:Accession: S25850

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 119-172 <BU2>

A:Cross-references: EMBL:X60055; NID:g396688; PIDN:CAA42653.1; PID:g396689

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1991

A:Accession: S36124

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 173-322, 'ELPT', 327-412 <BU3>

A:Cross-references: EMBL:X60091

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

R:Jakowlew, S.B.; Lechleider, R.; Geisler, A.G.; Kim, S.J.; Santa-Coloma, T.A.; Cubert, J.

Mol. Endocrinol. 6, 1285-1298, 1992

A:Title: Identification and characterization of the chicken transforming growth factor-h

A:Reference number: I51181; MUID:93024487; PMID:1406706

A:Accession: I51181

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-117 <JA2>

A:Cross-references: GB:S46000; NID:g257172; PIDN:AAB23575.1; PID:g257173

C:Genetics:

A:introns: 216/1; 252/1; 309/2; 363/3

A:Note: list of introns may be incomplete

C:Superfamily: inhibin

C:Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-300/Domain: propeptide #status predicted <PRO>

F:301-412/product: transforming growth factor beta-3 #status predicted <MAT>

F:74,142/Binding site: carbohydrate (Asn) #status predicted

Query Match 40.4%; Score 871.5; DB 2; Length 412;

Best Local Similarity 46.2%; Pred. No. 2.6e-62;

Matches 198; Conservative 56; Mismatches 114; Indels 61; Gaps 14;

QY 15 LVLVLTVTPGPAAGLSTCKTIDMELVKRRIEAIKQILSKLASPPSGQGVPPGPLP 74

Db 9 LVLSLSLFAVSLALSSCTLDLEHKRKRVEAIRQILSKLRLTSPPE--SVGPAHP 66

QY 75 BAVIALYNSTRDRVAGSVEPEPE-----PEADYAKVEYTRVLMVE--SGNQIYDKF 124

Db 67 YQILALYNSTRLL--BEMEBEKESQENTSEYAKETHKPKDMTQGLPEHNLGICP 124

QY 125 KGTPHSLYMLFNTSELREAVPEPVLSSRAELRLRL---KLKVEQHVELYQKYSND--- 177

Db 125 KGVTSNVFR-FNVS---SAEKNSTLFLRAEFLVRLVNPSPSSKRSEORIELFOILRPDEHI 180

QY 178 -SWRYLSNRLAPSDSPWLSFDVTGVVQWQTRREAIEGFRLSAH-----SSDSKD 229

Db 181 AKORYLSGRNVOTRGSPWLSFDVTGVREMLLHRESNLGLAISHCPCHTFQPNGDILE 240

QY 230 N---TLHVEINGFNSG---RRGLATI---HGMNRPFLMATPLERAOH--LHSSRRH 278

Db 241 NLHVELEIKFGIDSDDDYGRDGLRKLKKQKDLCHNPHEILMLPPHRLSEPTLGQQRKR 300

QY 279 ALDTSNYPDYVDYASLALDNTYCFSSTEKNCCVQOLYIDFRKDLGKMWIHEPKGYHANF 338

Db 301 -----ALDNTYCFNLEENCVRPLYIDFRQDLGKMWIHEPKGYFANF 343

QY 339 CLGCPYIWSLDTQYSKVLYLNQHNPGASAAAPCCVQALEPLPIVYVYVGRKPKVEQLSN 398

Db 344 CSQCPYLSADTTHSTVLGLYNILNPEASAPCCVQDLEPLTILYVYVGRTPKVEQLSN 403

QY 399 MIVRSCKCS 407

Db 404 MYVRSCKCS 412

RESULT 12

B61036

transforming growth factor beta-5 precursor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 16-Jul-1999

C:Accession: A34929; B61036

R:Kondaiah, P.; Sands, M.J.; Smith, J.M.; Fields, A.; Roberts, A.B.; Sporn, M.B.; Mel

J. Biol. Chem. 265, 1089-1093, 1990

A:Title: Identification of a novel transforming growth factor-beta (TGF-beta5) mRNA i

A:Reference number: A34929; MUID:90110090; PMID:2295601

A:Accession: A34929

A:Molecule type: mRNA

A:Residues: 1-382 <KON>

A:Cross-references: GB:J05180; NID:g214821; PIDN:AAA49968.1; PID:g214822

R:Roberts, A.B.; Rosa, F.; Roche, N.S.; Coligan, J.E.; Garfield, M.; Rebbert, M.L.; K

Growth Factors 2, 135-147, 1990

A:Title: Isolation and characterization of TGF-beta2 and TGF-beta5 from medium condit

A:Reference number: A61036; MUID:90253806; PMID:2340184

A:Accession: B61036

A:Molecule type: protein

A:Residues: 271-276, 'X', 278-284, 'XX', 287-299 <ROB>

C:Superfamily: growth factor

C:Keywords: transforming growth factor beta-5 #status experimental <MAT>

F:271-382/Product: transforming growth factor beta-5

Query Match 40.2%; Score 868.5; DB 2; Length 382;

Best Local Similarity 45.9%; Pred. No. 4e-62;

Matches 190; Conservative 55; Mismatches 122; Indels 47; Gaps 11;

QY 9 LPLLPLMLLVLTGCPAAGLSTCKTIDMELVKRRIEAIKQILSKLASPPSGQGV 68

Db 1 MEVLMMLLVLLVHLSSLSLSTCKRAVDMEYVKRRIEAIKQILSKLKDTPDV-DS 59

QY 69 PPGLPEAVLYLNSTRDRVAGSVEPE-----PEPEADYAKVEYTRVLMVESGNIYDK 123

Db 60 EKMTPVSEAFIYNSTLEIVREKATREEHVGHQDNIQDYAKQVYR---FESITELD- 115

QY 124 FKGTPLSLYMLFNTSELREAVPEPVLSSRAELRLRLKL--KYEQHVELYQKYSNDSW-- 179

Db 116 -----HEFKFKFNASHVRENVMNSLHHAELRMKYKQTDKNHDMELFQYQENGTH 170

QY 180 -RYLSNRLAPSDSPWLSFDVTGVVQWQTRREAIEGFRLSAH-----SSDSKDNLHVE 235

Db 171 SRYLESKYITPVTDDEWMSFDVTGVVQWQTRREAIEGFRLSAH-----SSDSKDNLHVE 226

QY 236 INGNFSRRGGLDATHGCM--NRPFLLMATPLERAOHLHSSRRHRLDNTNSYPDYVDY 293

```

Db      227 IEGF-PALRGDLASLSKENTKPYLMTSPAEIRDITVTSRKR----- 270
Qy      294 SLALDNYCYSSYKNCVQOLYIDFKDLGKWIHEPKGYHANFCLGCPPIYIWSIDTQY 353
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      271 --GVGGEYCFGNGPCVCKPLVINFKDLGKWIHEPKGYEANYCLGNCPIYIWSMDTQY 328
Qy      354 SKVLYALYNQHNPGASAPCCVQALPLPIYVYVGRKPKVEQLSNMIVRSCKCS 407
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      329 SKVLSLYNQHNPGASISPCVQDLPLPIYVYVGRKPKVEQLSNMIVRSNCSCS 382

RESULT 13
S01825
transforming growth factor beta-3 precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C:Accession: S01825
R:Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Mas
EMBO J. 7, 3737-3743, 1988
A:Title: A new type of transforming growth factor-beta, TGF-beta3.
A:Reference number: S01824; MUID:89091120; PMID:3208746
A:Accession: S01825
A:Molecule type: mRNA
A:Residues: 1-409 <DER>
A:Cross-references: EMBL:X14150; NID:g2127; PIDN:CAA32363.1; PID:g2128
C:Superfamily: inhibin
C:Keywords: growth factor
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-297/Domain: propeptide #status predicted <PRO>
F:298-409/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match      38.68; Score 834; DB 2; Length 409;
Best Local Similarity 44.68; Pred. No. 2.7e-59;
Matches 190; Conservative 58; Mismatches 122; Indels 56; Gaps 14;

Qy      15 LLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIKQILSKRLASPPSGQVPPGPLP 74
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      7 LVLLALLNFATVLSMSTCTTDFDHKKRVEAIRGQILSKRLTSPDPMSL--ANIP 64

Qy      75 EAVLYALYNSTR---DRVAGESVE--PEPEADYIAKEVTRVLV---ESGNQIYDKFG 126
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      65 TVQLDLYNSTRELLEEVHGERGDDCTQENTSEYIAKEIKFDMIOGLEHNDLAVCPKG 124

Qy      127 TPHSYLMFNTSELREAVEPVLISRAELRLRL---KLKVEQHVLYOKYSND---S 178
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      125 ITSIFR-FNVSSVEK---NETNLFRAEFVRLVNPFPNSSKRSQRIELFQILQDEHIAK 180

Qy      179 WRYLSNRLAPSDSPWLSFDVTGVVROWLTRREAEIGFRLSAH-----SSSDSKDN- 230
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      181 QRYIDGKNLPTRGAAEWLSFDVTDVREMLLRRESNLGLEISIHCPCHTFOPNGDILENI 240

Qy      231 --TLHVEINGFNS--GRRGDLATIHGM--NRPFLLLMATPLERAQH--LHSSRRALD 281
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      241 QEVMEIKFGVDSEDDFGDGLRKLKKEKSHPLILMIPDLDNPLGLGAQKRK--- 297

Qy      282 TNSYPYDVPDYASALDITNYCFSTKNCVQVLYIDFKDLGKWIHEPKGYHANFCLG 341
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      298 -----ALDTHNCFRNLEENCVRPLIYIDFRDLGKWKWHEPKGYANFCSG 343

Qy      342 PCPIYIWSIDTQYKVLALYNQHNPGASAPCCVQALPLPIYVYVGRKPKVEQLSNMIV 401
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      344 PCPYLRSDATTHSSVLGLYNLNPFEASAPCCVQDLPLPIYVYVGRKPKVEQLSNMIV 403

Qy      402 RSCKCS 407
      :   :   :   :   :
Db      404 KSCKCS 409

RESULT 14
A41397
transforming growth factor beta-3 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 16-Jul-1999

```

```

C:Accession: A41397; A61039; A61225
R:Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.
Mol. Endocrinol. 3, 1926-1934, 1989
A:Title: Complementary DNA cloning of the murine transforming growth factor-beta3 (TG
t tissues.
A:Reference number: A41397; MUID:90190650; PMID:2628730
A:Accession: A41397
A:Molecule type: mRNA
A:Residues: 1-410 <MIL>
A:Cross-references: GR:M32745; NID:g201949; PIDN:AAA40422.1; PID:g201950
R:Denhez, F.; Iafatis, R.; Kondalah, P.; Roberts, A.B.; Sporn, M.B.
Growth Factors 3, 139-146, 1990
A:Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mtTGF-beta3.
A:Reference number: A61039; MUID:91000714; PMID:2206556
A:Accession: A61039
A:Molecule type: mRNA
A:Residues: 1-410 <DEN>
R:Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.
Cell Growth Differ. 2, 77-83, 1991
A:Title: Cell lineage specificity of expression of the murine transforming growth fac
A:Reference number: A61225; MUID:91299576; PMID:2069871
A:Accession: A61225
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 285-410 <WAT>
C:Superfamily: inhibin
C:Keywords: glycoprotein; growth factor; growth regulation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-298/Domain: propeptide #status predicted <PRO>
F:299-410/Product: cell attachment (R-G-D) motif
F:299-410/Product: transforming growth factor beta-3 #status predicted <MAT>
F:72,133,140/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match      38.38; Score 827.5; DB 2; Length 410;
Best Local Similarity 44.28; Pred. No. 8.9e-59;
Matches 188; Conservative 57; Mismatches 127; Indels 53; Gaps 13;

Qy      15 LLMLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIKQILSKRLASPPSGQVPPGPLP 74
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      7 LVLLALLNLATISLSLSTCTTDFGHKKRVPKRAIKQILSKRLTSPPEPSVMT--HVP 64

Qy      75 EAVLYALYNSTR---DRVAGESVE--PEPEADYIAKEVTRVLV---SGNOIYDKFKG 126
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      65 YQVLYALYNSTRELLEEMHGEREGGCTQETSESEYIAKEIKHFKDMIOGLEAHNELAVCPKG 124

Qy      127 TPHSYLMFNTSELREAVEPVLISRAELRLRL---KLKVEQHVLYOKYSND---S 178
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      125 ITSIVFR-FNVSSVEK---NGTNLFRAEFVRLVNPFPSSKRTQRIELFQILRPDEHIAK 180

Qy      179 WRYLSNRLAPSDSPWLSFDVTGVVROWLTRREATEGFRLSAH-----SSSDSKDN- 230
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      181 QRYIGGKNLPTRGTAEWLSFDVTDVREMLLRRESNLGLEISIHCPCHTFOPNGDILENV 240

Qy      231 --TLHVEINGFNS--GRRGDLATIHGM--NRPFLLLMATPLERAQHLHSSRRALD 282
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      241 HEVMEIKFGVDNEDDHGRGDLKRLKQKQDHNPHILMIPPHRLDSDPSQGSQK--- 296

Qy      283 NSYPYDVPDYASALDITNYCFSTKNCVQVLYIDFKDLGKWKWHEPKGYHANFCLG 342
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      297 -----KRALDITNYCFRNLEENCVRPLIYIDFRDLGKWKWHEPKGYANFCSG 345

Qy      343 PCPIYIWSIDTQYKVLALYNQHNPGASAPCCVQALPLPIYVYVGRKPKVEQLSNMIV 402
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      346 CPYLRSDATTHSTVGLIYNLNPFEASAPCCVQDLPLPIYVYVGRKPKVEQLSNMIV 405

Qy      403 RSCKCS 407
      :   :   :   :   :
Db      406 KSCKCS 410

RESULT 15
A55706
transforming growth factor beta-3 precursor - rat

```

C:Species: Rattus norvegicus (Norway rat)  
C:Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 16-Jul-1999  
C:Accession: A55706; B40699; S36042  
R:Wang, J.; Kulliszewski, M.; Yee, W.; Sedlackova, L.; Xu, J.; Tseu, I.; Post, M.  
J. Biol. Chem. 270, 2722-2728, 1995  
A:Title: Cloning and expression of glucocorticoid-induced genes in fetal rat lung fibroblasts  
A:Reference number: A55706; MUID:95155340; PMID:7852342  
A:Accession: A55706  
A:Molecule type: mRNA  
A:Residues: 1-410 <WAS>  
A:Cross-references: GB:U03491  
A:Note: it is uncertain whether Met-1 is the initiator  
R:McKinnon, R.D.; Piras, G.; Ida Jr., J.A.; Dubois-Dalcq, M.  
J. Cell Biol. 121, 1397-1407, 1993  
A:Title: A role for TGF-beta in oligodendrocyte differentiation.  
A:Reference number: A40699; MUID:53286190; PMID:8509457  
A:Accession: B40699  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 157-211 <MCK>  
A:Cross-references: EMBL:X71903; NID:g311326; PIDN:CAA50722.1; PID:g311327  
C:Superfamily: inhibin

Query Match.	38.38;	Score	826.5;	DB	2;	Length	410;		
Best local Similarity	44.28;	Pred.	No. 1.le-58;						
Matches	188;	Conservative	57;	Mismatches	127;	Indels	53;	Gaps	13;
Qy	15	LLMLLVITPCRPAAGLSTCTIDMELVKRRIEAI	RGQILSKL	SLAPPSGDVPPG	PLP	74			
Db	7	LVVLLALLNATVSLSTCTIDFGHIKKRVEAIR	QILSKL	RLTSPPEPSVMT		-HVP	64		
Qy	75	EAUVALYNSTR---	DRVAGKSVE--	PPPEPADYAYAKEVTRVL	VMVE---	SGNQIYDFKPG	126		
Db	65	YOVVALYNSTRLEEL	EEHMGHREEGCTQETSE	SEYYAKEIHKFDMIOGLA	EHNELAVCPKG	124			
Qy	127	TPHSLYMLFNTSEL	REAVPVLVLSRAELRLRL---	KLKVEGHVELYOKYSND---	S	178			
Db	125	ITSKVPR-FNVSSVEK---	NGTNLFRAEFRLV	VPNPNSSKTEQRIEL	FQLRLRDEHIAK	180			
Qy	179	WRYLSNKLAPSDS	PEWLSLTDVTVGVVROWITR	REAI	EGFRLSAH-----	SSSDSKDN-	230		
Db	181	QRYIGGKNLPTRGTA	EAWLSLTDVTDVREWLRL	RRESNLGLEISIHCPCHIT	FOPNGDILENV	240			
Qy	231	-TLHVEINGFNS-	-GRKDLATIHGM--	NRPFLLIMATPLERAQHLHSSRR	HALDIT	282			
Db	241	HEVMEIRFKGVND	EDDGRGLGRKKQKQKOHNNPHLIM	IPPHRLSDSPGGOGORK	----	296			
Qy	283	NSPYDPDYASL	ALDNTNYCFSSTEKNCV	RQLYIDFRKDLGKWKI	HEPKGYHANFLGP	342			
Db	297	-----	KRALDNTYCFRNLEEN	CCVRPLYIDFRQDLGKWKV	HEPKGYANFCSGP	345			
Qy	343	CPYIWSLDTQY	SKVALYNCHNP	CASAAAPCCV	CPQALEPLTVI	YVYGRKPKVEQLSNM	TVR	402	
Db	346	CPYLRSDTHST	VLGYNTYNLP	NEASASPC	CCVPQDLEPLT	IYVYGRTPKVEQLSNM	VYVK	405	
Qy	403	SCKCS	407						
Db	406	SCKCS	410						

Search completed: October 7, 2003, 18:05:04  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:52:07 ; Search time 23 Seconds  
(without alignments)  
832.168 Million cell updates/sec

Title: US-10-017-372E-37

Perfect score: 2159  
Sequence: 1 MAPSGRLRLPLLLPLLLLV.....GRKPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2030.5	94.0	390	1	TGFL_PIG
2	1929.5	89.4	390	1	P07200 sus scrofa
3	1916.5	88.8	390	1	P50414 ovis aries
4	1914.5	88.7	390	1	TGFL_HUMAN
5	1893.5	87.7	390	1	P09533 cercopithec
6	1893.5	85.2	390	1	P54831 canis famil
7	1835.5	85.0	390	1	O19011 equus cabal
8	1835.5	85.0	390	1	P04202 mus musculu
9	1826.5	84.6	390	1	P17246 rattus norv
10	1572.5	72.8	315	1	TGFL_CAVPO
11	1047	48.5	373	1	TGFL_BOVIN
12	868.5	40.2	382	1	TGFL_CHICK
13	852.5	39.5	412	1	TGFL_XENLA
14	834	38.6	409	1	TGFL_PIG
15	827.5	38.3	410	1	TGFL_MOUSE
16	826.5	38.3	412	1	TGFL_RAT
17	824.5	38.2	412	1	TGFL_HUMAN
18	824	38.2	414	1	TGFL_MOUSE
19	809	37.5	414	1	TGFL_HUMAN
20	808.5	37.4	413	1	TGFL_XENLA
21	808	37.4	412	1	TGFL_CHICK
22	805	37.3	435	1	TGFL_PIG
23	805	37.3	442	1	TGFL_RAT
24	482	22.3	112	1	TGFL_BOVIN
25	323.5	15.0	375	1	GDF8_MEIGA
26	321.5	14.9	375	1	GDF8_CHICK
27	311.5	14.4	375	1	GDF8_PIG
28	310.5	14.4	375	1	GDF8_PAPHA
29	308.5	14.3	376	1	GDF8_RAT
30	307.5	14.2	375	1	GDF8_HUMAN
31	306.5	14.2	376	1	GDF8_MOUSE
32	305.5	14.2	375	1	GDF8_SHEEP
33	300	13.9	405	1	GDFB_MOUSE

34	300	13.9	407	1	GDFB_HUMAN
35	296.5	13.7	375	1	GDF8_BOVIN
36	280.5	13.0	374	1	GDF8_BRARE
37	278.5	12.9	345	1	GDFB_RAT
38	275	12.7	425	1	IHBA_SHEEP
39	272.5	12.6	424	1	IHBA_PIG
40	270.5	12.5	424	1	IHBA_MOUSE
41	270.5	12.5	424	1	IHBA_RAT
42	270	12.5	425	1	IHBA_BOVIN
43	269.5	12.5	426	1	IHBA_HORSE
44	268.5	12.4	426	1	IHBA_HUMAN
45	265.5	12.3	424	1	IHBA_CHICK

ALIGNMENTS

RESULT 1  
TGFL\_PIG  
ID TGFL\_PIG STANDARD; PRT; 390 AA.  
AC P07200; P08832;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
GN TGFB1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=87174844; PubMed=3470708;  
RA Derynck R., Rhee L.;  
RT "Sequence of the porcine transforming growth factor-beta precursor.";  
RL Nucleic Acids Res. 15:3187-3187(1987).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
RC STRAIN=Miniature swine;  
RX MEDLINE=89054010; PubMed=2461367;  
RA Kondalah P., van Obberghen-Schilling E., Ludwig R.L., Dhar R.,  
RT Sporn M.B., Roberts A.B.;  
RT "cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
RT Evidence for alternate splicing and polyadenylation.";  
RN J. Biol. Chem. 263:18313-18317(1988).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
RX MEDLINE=88335639; PubMed=3166520;  
RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;  
RT "Nucleotide sequence of chicken transforming growth factor-beta 1  
RT (TGF-beta 1).";  
RL Nucleic Acids Res. 16:8730-8730(1988).  
RN [4]  
RP SHOWS THAT REF.3 SEQUENCE IS FROM PIG.  
RA Jakowlew S.B.;  
RL Unpublished observations (MAR-1996).  
RN [5]  
RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
RX Wimmers K., Chomdej S., Ponsuksili S., Schellander K.;  
RT "Polymorphism in the porcine transforming growth factor beta 1  
RT gene.";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 279-322.  
RX MEDLINE=87102890; PubMed=2879635;  
RA Cheifetz S., Weatherbee J.A., Tsang M.L.S., Anderson J.K., Mole J.E.,  
RA Lucas R., Massague J.;  
RT "The transforming growth factor-beta system, a complex pattern of  
RT cross-reactive ligands and receptors.";  
RL Cell 48:409-415(1987).  
CC -1- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL





```
DR SMART: SM00204; TGFβ: 1;
DR PROSITE: PS00250; TGFβ_BETA_1;
RW Growth factor; Mitogen; Glycoprotein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 278 POTENTIAL.
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
FT DISULFID 285 294 BY SIMILARITY.
FT DISULFID 293 356 BY SIMILARITY.
FT DISULFID 322 387 BY SIMILARITY.
FT DISULFID 326 389 BY SIMILARITY.
FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 82 82 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 390 AA: 44291 MW: 102472994840057 CRC64;

Query Match 89.4%; Score 1929.5; DB 1; Length 390;
Best Local Similarity 89.7%; Pred. No. 6.8e-145;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;

QY 1 MAPSGRLRLPLLLPLMLLVITGPRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLA 60
DB 1 MPSPGLRLPLLLPLMLLVITGPRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLA 60
QY 61 SPSPGQDVPVPGPLPAVALYNSTRDVRAGESVEPEPEADYAKETRVLMVSGNQI 120
DB 61 SPSPGQDVPVPGPLPAVALYNSTRDVRAGESVEPEPEADYAKETRVLMVSGNQI 120
QY 121 YDKFKGTPHSYMLFNTSELREAVPEPVLLSRAELRLRLKLKVEQHVLYOKYSNDSWR 180
DB 121 YDKMKSSHSYMFNTSELREAVPEPVLLSRAELRLRLKLKVEQHVLYOKYSNDSWR 180
QY 181 YLSNRLAPSDPEWLSFVTVGVVQWLTREAIKGFRLSAHSSDSKNTLHVEINGFN 240
DB 181 YLSNRLAPSDPEWLSFVTVGVVQWLTREAIKGFRLSAHSSDSKNTLHVEINGFN 240
QY 241 SGRRGLATIGHMNPFLLLMATPLERAQHLHSSRRRLDNTNSYPVDPVYASIALDTN 300
DB 241 SGRRGLATIGHMNPFLLLMATPLERAQHLHSSRRRLDNTNSYPVDPVYASIALDTN 300
QY 301 YCFSTENKCCVRQLYIDFRKDLGKWKTHPEKGYHANFCLGCPPIWISLDTOYSKVLALY 360
DB 284 YCFSTENKCCVRQLYIDFRKDLGKWKTHPEKGYHANFCLGCPPIWISLDTOYSKVLALY 343
QY 361 NOHNPAGSAAPCCVQPALEPLPIVYVGVKPKKVEQLSNMIVRSCKCS 407
DB 344 NOHNPAGSAAPCCVQPALEPLPIVYVGVKPKKVEQLSNMIVRSCKCS 390

RESULT 3
TGFβ_HUMAN
ID TGFβ_HUMAN STANDARD; PRT; 390 AA.
AC P01137; Q9UCG4;
DT 21-JUL-1986 (Rel. 01, Created)
DE 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1).
GN TGFβ1 OR TGFβ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87174845; PubMed=3470709;
RA Derynck R., Rhee L., Chen E.Y., van Tilburg A.;
RT "Intron-exon structure of the human transforming growth factor-beta precursor gene.";
RL Nucleic Acids Res. 15:3188-3189(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85296301; PubMed=3861940;
RA Derynck R., Roberts A.B., Sporn M.B., Gieddel D.V.;
RT "Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells.";
RL Nature 316:701-705(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse J.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner I., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Biakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 279-390 FROM N.A.
RX TISSUE=Carcinoma;
RA Urushizaki Y., Niitsu Y., Terui T., Koshida Y., Mahara K., Kohgo Y.,
RA Urushizaki I., Takahashi Y., Ito H.;
RT "Cloning and expression of the gene for human transforming growth factor-beta in Escherichia coli.";
RL Tumor Res. 22:41-55(1987).
RN [5]
RP SEQUENCE OF 279-329.
RX TISSUE=Bladder carcinoma;
RA Bourdrel L., Lin C.-H., Lauren S.L., Elmore R.H., Sugarman B.J.,
RA Hu S., Westcott K.R.;
RT "Recombinant human transforming growth factor-beta 1: expression by Chinese hamster ovary cells, isolation, and characterization.";
RL protein Expr. Purif. 4:130-140(1993).
RN [6]
RP SEQUENCE OF 279-301.
RX MEDLINE=85131019; PubMed=2982829;
RA Massague J., Like B.;
RT "Cellular receptors for type beta transforming growth factor. Ligand binding and affinity labeling in human and rodent cell lines.";
RL J. Biol. Chem. 260:2636-2645(1985).
RN [7]
RP STRUCTURE BY NMR OF 279-390.
RX MEDLINE=93144319; PubMed=842942;
RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,
RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
RA Torchia D.A.;
RT "Transforming growth factor beta 1: NMR signal assignments of the recombinant protein expressed and isotopically enriched using Chinese hamster ovary cells.";
RL Biochemistry 32:1152-1163(1993).
RN [8]
RP STRUCTURE BY NMR OF 279-390.
RX MEDLINE=93144320; PubMed=842943;
RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,
RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
RA Torchia D.A.;
RT "Transforming growth factor beta 1: secondary structure as determined by heteronuclear magnetic resonance spectroscopy.";
RL Biochemistry 32:1164-1171(1993).
RN [9]
```



```

CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16658; AAA35369.1; -.
CC PIR: A26960; A26960.
CC HSP: P01137; IKLA.
CC InterPro: IPR002400; GF_cysknot.
CC InterPro: IPR003911; TGF_TGFB.
CC InterPro: IPR001839; TGFB.
CC InterPro: IPR001111; TGFB_N.
CC Pfam: PF00019; TGF-beta; 1.
CC Pfam: PF00688; TGFB_propeptide; 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC ProDom: PD000357; TGFB; 1.
CC SMART: SM00204; TGFB; 1.
CC PROSITE: PS00250; TGF_BETA_1; 1.
CC Growth factor; Mitogen; Glycoprotein; Signal.
CC SIGNAL 16 POTENTIAL.
CC FT PROPEP 17 278
CC FT CHAIN 279 390
CC FT DISULFID 285 294
CC FT DISULFID 293 356
CC FT DISULFID 322 387
CC FT DISULFID 326 389
CC FT DISULFID 355 355
CC FT CARBOHYD 82 82
CC FT CARBOHYD 136 136
CC FT CARBOHYD 176 176
CC FT SITE 244 246
CC SEQUENCE 390 AA; 44356 MW; DPF63E2BBA4320E CRC64;
CC -----
Query Match 88.7%; Score 1914.5; DB 1; Length 390;
Best Local Similarity 89.4%; Pred. No. 1e-143;
Matches 364; Conservative 10; Mismatches 16; Indels 17; Gaps 1;
Oy 1 MAPSGRLRLPLLLPLLLVLTGPPAAGLSTCKTIDMEVAKRRIEAIKQILSKRLA 60
Db 1 MPSPGLRLPLLLPLLLVLTGPPAAGLSTCKTIDMEVAKRRIEAIKQILSKRLA 60
Oy 61 SPPSGDVPGLPPLPAVLALYNSTRDRVAGSEPEPEADYAKVETRVLMVESGNOI 120
Db 61 SPPSGDVPGLPPLPAVLALYNSTRDRVAGSEPEPEADYAKVETRVLMVETHNEI 120
Oy 121 YDKFKGTPHLYMFLNFTSELREAVPEPVLLSRAELRLRLKLVQHVLYKYSNDSWR 180
Db 121 YDKFKGTPHLYMFLNFTSELREAVPEPVLLSRAELRLRLKLVQHVLYKYSNDSWR 180
Oy 181 YLSNRLAPSPEWLSFDVTVGVVROWLTREATEGFRLSAHSSDSKNTLHVEINGFN 240
Db 181 YLSNRLAPSPEWLSFDVTVGVVROWLTREATEGFRLSAHSSDSKNTLQVDINGFT 240
Oy 241 SGRRGLDLAGHGNRPFLLMATPLERAQHLSRRHRLDNTNSPYDVPDVSALALDTN 300
Db 241 TGRGDLATIGHGNRPFLLMATPLERAQHLSRRHRLDNTNSPYDVPDVSALALDTN 300
Oy 301 YCFSTEKNCCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTOYSKVLALY 360
Db 284 YCFSTEKNCCVQRLYIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTOYSKVLALY 343
Oy 361 NQHNFGASAAPCCVPOALEPLPIVTVYVGRKPKVEQLSNMIVRSCKS 407
Db 344 NQHNFGASAAPCCVPOALEPLPIVTVYVGRKPKVEQLSNMIVRSCKS 390

```

```

RESULT 5
TGFI_CANFA STANDARD; PRT; 390 AA.
ID TGFI_CANFA
AC P54831;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1).
OS Homo sapiens.
GN TGFBI.
OC Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Jugular vein endothelial;
RX MEDLINE=95237630; PubMed=7721110;
RA Manning A.M., Auchampach J.A., Drong R.F., Slightom J.L.;
RT "Cloning of a canine cDNA homologous to the human transforming growth
RI factor-beta 1-encoding gene.";
RL Gene 155:307-308(1995).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L34956; AAA51458.1; -.
CC PIR: JC4023; JC4023.
CC HSP: P01137; IKLA.
CC InterPro: IPR002400; GF_cysknot.
CC InterPro: IPR003911; TGF_TGFB.
CC InterPro: IPR001839; TGFB.
CC InterPro: IPR001111; TGFB_N.
CC Pfam: PF00019; TGF-beta; 1.
CC Pfam: PF00688; TGFB_propeptide; 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC ProDom: PD000357; TGFB; 1.
CC SMART: SM00204; TGFB; 1.
CC PROSITE: PS00250; TGF_BETA_1; 1.
CC Growth factor; Mitogen; Glycoprotein; Signal.
CC SIGNAL 23 POTENTIAL.
CC FT PROPEP 24 278
CC FT CHAIN 279 390
CC FT DISULFID 285 294
CC FT DISULFID 293 356
CC FT DISULFID 322 387
CC FT DISULFID 326 389
CC FT DISULFID 355 355
CC FT CARBOHYD 82 82
CC FT CARBOHYD 136 136
CC FT CARBOHYD 176 176
CC FT SITE 244 246
CC SEQUENCE 390 AA; 44185 MW; EB4780E88B7B590E CRC64;
CC -----
Query Match 87.7%; Score 1893.5; DB 1; Length 390;
Best Local Similarity 88.7%; Pred. No. 4.7e-142;
Matches 361; Conservative 12; Mismatches 17; Indels 17; Gaps 1;

```





```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52498; CAA36741.1; -.
CC PIR: S10219; S10219.
CC HSSP: P01137; IKLA.
CC InterPro: IPR002400; GF_cysknot.
CC InterPro: IPR003911; TGF_TGFB.
CC InterPro: IPR001839; TGFb.
CC InterPro: IPR001111; TGFb_N.
CC Pfam: PF00019; TGF-beta; 1.
CC Pfam: PF00688; TGFb-propeptide; 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC PRINTS: PR01423; TGFbETA.
CC ProDom: PD000357; TGFb; 1.
CC SMART: SM00204; TGFb; 1.
CC PROSITE: PS00250; TGF_BETA_1; 1.
CC Growth factor; Mitogen; Glycoprotein; Signal.
KW SIGNAL 1 23
FT PROPEP 24 278
FT CHAIN 279 390
FT DISULFID 285 294
FT DISULFID 293 356
FT DISULFID 322 387
FT DISULFID 326 389
FT DISULFID 355 355
FT CARBOHYD 82 82
FT CARBOHYD 136 136
FT CARBOHYD 176 176
FT SITE 244 246
SQ SEQUENCE 390 AA; 44329 MW; 5E21108ED50D853C CRC64;
Query Match 85.0%; Score 1835.5; DB 1; Length 390;
Best local Similarity 85.0%; Pred. No. 1 8e-137;
Matches 346; Conservative 14; Mismatches 30; Indels 17; Gaps 1;
QY 1 MAPSGRLRLPLLLPLWLLVLTGPRPAAGLSTCKTIDMELVKRKRTEARIGQILSKRLRA 60
Db 1 MPSPGLRLPLLLPLWLLVLTGPRPAAGLSTCKTIDMELVKRKRTEARIGQILSKRLRA 60
QY 61 SPSPQSDVPGPLPEAVLA.YNSTRDRVAGESVEPEPEADYAKETRVLMVSGNOI 120
Db 61 SPSPQSDVPGPLPEAVLA.YNSTRDRVAGESADPEPEADYAKETRVLMVDRNNAI 120
QY 121 YDKFKGTPHSLYMLFTISE.REAVPEPVLSSRAELRLKLKVEQHVLYQKYSNDSWR 180
Db 121 YDKTKDTHSIYMFNTSDIREAVPEPPLLSRAELRLQRFKSTVEQHVLYQKYSNNSWR 180
QY 181 YLSNRLAPSDSPWLSEFD.TGVVROWLRTREAEGRFLSAHSSDSKONTLHVEINGEN 240
Db 181 YLGNRLLTPTDPEWLSFD.TGVVROWLNGDGIQGFRESAHSCSDSKDNLVHVEINGIS 240
QY 241 SGRRGDLATHGCMNRPELL.MATPLERAQHLSSRRRALDNTSYDPDVAASLALDNT 300
Db 241 PKRGDLGTHDMNRPELL.MATPLERAQHLSSRRR-----ALDNT 283
QY 301 YCFSTKNCVQRLYIDF.KDGLGWKWIHEPKGYHANFCLGPCPYTWSLDTQYSKVLALY 360
Db 284 YCFSTKNCVQRLYIDF.KDGLGWKWIHEPKGYHANFCLGPCPYTWSLDTQYSKVLALY 343
QY 361 NQHPGASAPCCVPALEPLPIYVYVGRKPKVEQLSNMIVRSCKS 407
Db 344 NQHPGASAPCCVPALEPLPIYVYVGRKPKVEQLSNMIVRSCKS 390
RESULT 9
TGFL_CAVPO
ID TGFL_CAVPO STANDARD; PRT; 390 AA.
AC Q9Z1Y6; Q9ZB3; Q9R148;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1).
GN TGFBI.
```

```

FT CONFLICT 286 286 F -> S (IN REF. 2).
FT CONFLICT 309 309 K -> E (IN REF. 2).
FT CONFLICT 322 322 C -> R (IN REF. 2).
FT CONFLICT 350 350 A -> G (IN REF. 2).
SQ SEQUENCE 390 AA: 44328 MW: 1539F849BA0C0FF1 CRC64;

Query Match 84.6%; Score 1826.5; DB 1; Length 390;
Best Local Similarity 84.8%; Pred. No. 9.2e-137; Indels 17; Gaps 1;
Matches 345; Conservative 15; Mismatches 30;

Qy 1 MAPSGRLPLPLLLPLMLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKLRLA 60
Db 1 MPPSRLLPLPLLLPLMLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKQILSKLRLA 60
Qy 61 SPPSQGVPPGCPLEAVLALYNSTRDVAGSVPEPEPEADYAKVETRLVMVESGNOI 120
Db 61 SPPSQGVPPGCPLEAVLALYNSTRDVAGSVPEPEPEADYAKVETRLVMVDSHNI 120
Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVQHVLYKYSNDSWR 180
Db 121 YKSIETVAHSIYMFNTSELREAVPDPLLSRAELRMQRLNVLQHVLYKYSNNSWR 180
Qy 181 YLSNKLAPSDSEPLSFVDTGVVQWLTTRAELEGFRLSAHSSDSKDNLTLHVEINGFN 240
Db 181 YLSNQLTSPDTEWLSFVDTGVVQWLSQCEEGFRLSAHSCDSKDNLTLRVEINGIG 240
Qy 241 SGRGOLATIHGMNRPFLMLMATPLERAQHLHSSRRHALDTNSYPYDVPDYASALDTN 300
Db 241 PKRGDLAAIHGMNRPFLMLMATPLERAQHLHSSRRHR-----GLDTN 283
Qy 301 YCFSTERNKCCVRLYIDFRDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 360
Db 284 YCFSTERNKCCVRLYIDFRDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALY 343
Qy 361 NOHNPASAAPCCVQALPELPYIVYVGRKPKVEQLSNMIVRSCKS 407
Db 344 NOHNPASAAPCCVQALPELPYIVYVGRKPKVEQLSNMIVRSCKS 390

RESULT 10
TGFL_BOVIN STANDARD: PRT: 315 AA.
AC P18341;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1) (Fragment).
GN TGFB1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91042552; PubMed=313459;
RA van Oberghe-Schilling E., Kondalah P., Ludwig R.L., Sporn M.B.,
RA Baker C.C.;
RT "Complementary deoxyribonucleic acid cloning of bovine transforming
RT growth factor-beta 1."
RL Mol. Endocrinol. 1:693-698(1987).
RN [2]
RP SUBUNITS.
RC TISSUE=Bone;
RX MEDLINE=92129307; PubMed=1733936;
RA Ogawa Y., Schmidt D.K., Dasch J.R., Chang R.J., Glaser C.B.;
RT "Purification and characterization of transforming growth factor-beta
RT 2.3 and -beta 1.2 heterodimers from bovine bone."
RL J. Biol. Chem. 267:2325-2328(1992).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE

```

```

CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: Homodimer; disulfide-linked. Heterodimers of TGF-beta 1/2
CC have been found in bone.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M36271; AAA30778.1; .
CC PIR: A40057; A40057.
CC HSP: P01137; IKLA.
CC InterPro: IPR002400; GF_cysknot.
CC InterPro: IPR003911; TGF_TGFB.
CC InterPro: IPR001839; TGFB.
CC InterPro: IPR001111; TGFB_N.
CC Pfam: PF00019; TGF-beta; 1.
CC PRINTS: PR00688; TGFB_propeptide; 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC PRINTS: PR01423; TGFBETA.
CC PRODOM: PD000357; TGFB; 1.
CC SMART: SM00204; TGFB; 1.
CC PROSITE: PS00250; TGF_BETA_1; 1.
CC Growth factor: Mitogen; Glycoprotein.
CC FT NON_TER 1 1
CC FT PROPEP <1 203
CC FT CHAIN 204 315 TRANSFORMING GROWTH FACTOR BETA 1.
CC FT DISULFID 210 219 BY SIMILARITY.
CC FT DISULFID 218 281 BY SIMILARITY.
CC FT DISULFID 247 312 BY SIMILARITY.
CC FT DISULFID 251 314 BY SIMILARITY.
CC FT DISULFID 280 280 INTERCHAIN (BY SIMILARITY).
CC FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT SITE 169 171 CELL ATTACHMENT SITE (POTENTIAL).
CC SQ SEQUENCE 315 AA: 36269 MW: C2717A23D94E00E CRC64;

Query Match 72.8%; Score 1572.5; DB 1; Length 315;
Best Local Similarity 88.9%; Pred. No. 7.6e-117; Indels 17; Gaps 1;
Matches 295; Conservative 9; Mismatches 11;

Qy 76 AVLALYNSTRDVAGSVPEPEADYAKVETRLVMVESGNOIYDKFKGTPHSLYMLF 135
Db 1 AILALYNSTRDVAGSAETEPEADYAKVETRLVMVEYGNKIYDKMKSSHSITMFF 60
Qy 136 NTSLEAVPEPVLLSRAELRLRLKLVQHVLYKYSNDSWRYLSNKLAPSDSPW 195
Db 61 NTSLEAVPEPVLLSRADVRLRLKLVQHVLYKYSNNSWRYLSNKLAPSDSPW 120
Qy 196 LSFVDTGVVQWLTTRAELEGFRLSAHSSDSKDNLTLHVEINGFNSGRGDLATIHGMNR 255
Db 121 LSFVDTGVVQWLTTRAELEGFRLSAHSCDSKDNLTQVDINGFSSGRGDLATIHGMNR 180
Qy 256 PFLLMATPLERAQHLHSSRRHALDTNSYPYDVPDYASALDTNYCFSTERNKCCVRL 315
Db 181 PFLLMATPLERAQHLHSSRRH-----ALDTNYCFSTERNKCCVRL 223
Qy 316 YIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALYNQHNPGASAPCCVP 375
Db 224 YIDFRKDLGKWKIHEPKGYHANFCLGCPYIWSLDTQYSKVLALYNQHNPGASAPCCVP 283
Qy 376 QALEPLPIVYVGRKPKVEQLSNMIVRSCKS 407
Db 284 QALEPLPIVYVGRKPKVEQLSNMIVRSCKS 315

```



FT	CARBOHYD	153	153	N-LINKED (GLCNAC, ...)	(POTENTIAL).
SITE		224	226	CELL ATTACHMENT SITE	(POTENTIAL).
SQ	SEQUENCE	373 AA;	42634 MW;	99033479C8552E5 CRC64;	
	Query Match		48.5%;	Score 1047;	DB 1; Length 373;
	Best Local Similarity		52.9%;	Pred. No. 2.9e-75;	
	Matches	207;	Conservative	52;	Mismatches 100; Indels 32; Gaps 8;
QY	30	LS	CTK	TIDMELVKRKRLEAIRGQILSKRLASPPSGDVPVPGPLPAVALYALYNSTRDVA	89
DB	2	LS	TCQ	LDLEAAKKRIEAVRGQILSKRLTAPPASETPPPRELPDQDVALYNSTQELLK	61
QY	90	GES	-V	PEPEADYAKAVTRVLMVESGNIQYDKFKGPHSLYMLFNTSELREAVPEPV	148
DB	62	OR	ARL	PPDPGDEYNAKELRRIPMETTWAGAMEHQPSHSIFFVFNVSRRRG-CRPT	120
QY	149	LL	SRA	ELRLRLKLK-----VEQHVELYQKYSNDSWRY-SNRLALASDSEPLSFDVTGV	203
DB	121	LL	HRA	ELRLRLQKAADASAGTEORLELYQGYGNASWRYLHGSRVRAADDEWLSFDVDA	180
QY	204	V	ROL	TRRAIEGRFSLSAHSSD---SKDNTLHVEINGFNSGRRGDLATIGHNMR--PFL	258
DB	181	V	HO	LSSGSELLGVFKLSVHCPCWGPCHAEEMRISTEGFEQ-ORGDMQSIATKXRRVPYV	239
QY	259	LL	MAT	PLERAQHLLHSRRHRRALDTSNPYDVPDYASLALDNTYCF--SSTEKKNCVRLQY	316
DB	240	L	AA	PAERANFLHSARRR-----DLSDTYCFGPGTDEKNCVRLPY	282
QY	317	ID	RK	DLGNKWIHEPKGYHANFCLGHCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQ	376
DB	283	ID	RK	DLQWNIHEPKGYMANFCMCPYIWSADTQITKVLALYNQHNPGASAAPCCVPQ	342
QY	377	AL	PL	PTIYVYVGRKPKVEQLSNMIVRSCKCS	407
DB	343	T	LD	PLTIYVYVGRNVRVEQLSNMIVRACKCS	373

RESULT 12

TGFI_XENLA	ID	TGFI_XENLA	STANDARD;	PRT;	382 AA.
AC	AC	P16176;			
DT	DT	01-APR-1990 (Rel. 14, Created)			
DT	DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	DE	Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 5).			
OS	OS	Xenopus laevis (African clawed frog).			
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
OC	OC	Xenopodinae; Xenopus.			
OX	OX	NCBI_TaxID=8355;			
RN	RN	[1]			
RP	RP	SEQUENCE FROM N.A.			
RX	RX	MEDLINE=90110090; PubMed=2295601;			
RA	RA	Kondaiah P., Sands M.J., Smith J.M., Fields A., Roberts A.B.,			
RA	RA	Sporn M.B., Melton D.A.;			
RT	RT	"Identification of a novel transforming growth factor-beta (TGF-beta			
RT	RT	5) mRNA in Xenopus laevis";			
RT	RT	J. Biol. Chem. 265:1089-1093(1990).			
RN	RN	[2]			
RP	RP	SEQUENCE FROM N.A.			
RA	RA	Vempati U.D., Kondaiah P.;			
RL	RL	Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.			
CC	CC	-!- FUNCTION: IMPORTANT ROLE IN CERTAIN ASPECTS OF DIFFERENTIATION.			
CC	CC	-!- SUBUNIT: Homodimer; disulfide-linked.			
CC	CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	CC	-!- SIMILARITY: Belongs to the TGF-beta family.			
CC	CC	-----			
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	CC	use by non-profit institutions as long as its content is in no way			
CC	CC	modified and this statement is not removed. Usage by and for commercial			
CC	CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).			

```
or send an email to license@lsb-sib.ch).
-----
CC CC EMBL; J05180; AAA49968.1; -.
CC CC EMBL; AF009335; AAB64441.1; -.
CC CC EMBL; AF009331; AAB64441.1; JOINED.
CC CC EMBL; AF009332; AAB64441.1; JOINED.
CC CC EMBL; AF009333; AAB64441.1; JOINED.
CC CC EMBL; AF009334; AAB64441.1; JOINED.
CC CC PIR; A34229; B61036.
CC CC HSSP; P01137; IKLA.
CC CC InterPro; IPR002400; GF_cysknot.
CC CC InterPro; IPR003911; TGF_TGFB.
CC CC InterPro; IPR001839; TGFB.
CC CC InterPro; IPR001111; TGFB_N.
CC CC Pfam; PF00019; TGF-beta; 1.
CC CC Pfam; PF00688; TGF-propeptide; 1.
CC CC PRINTS; PR00438; GFCYSKNOT.
CC CC PRINTS; PR01423; TGFBETA.
CC CC PRODOM; PD000357; TGFB; 1.
CC CC SMART; SM00204; TGFB; 1.
CC CC PROSITE; PS00250; TGF_BETA_1; 1.
CC CC Growth factor; Mitogen; Glycoprotein; Signal.
KW SIGNAL 1 21
FT CHAIN 22 270
FT PROPEP 271 382
FT CHAIN 271 382
FT DISULFID 277 286
FT DISULFID 285 348
FT DISULFID 314 379
FT DISULFID 318 381
FT DISULFID 347 381
FT CARBOHYD 73 73
FT CARBOHYD 123 123
FT CARBOHYD 166 166
FT SITE 234 236
SQ SEQUENCE 382 AA; 44200 MW; 1034621C917AAE15 CRC64;

Query Match
Best Local Similarity 40.2%; Score 868.5; DB 1; Length 382;
Matches 190; Conservative 55; Mismatches 122; Indels 47; Gaps 11;

Qy 9 LPLLPLMLLVLPGRPAAGLSCTCKIDMELVKKRIEIRGILSKRLASPPSGDV 68
Db 1 MEVLMLLVLLVLSLAMSLSCTKAVDMEEVKRRIRIEIRGILSKRLDKTPDV 59
Qy 69 PGGLPEAVLALYNSTRDEVAGESVEPE-----PEPEADYVAKVETVLMVESCNOIYDK 123
Db 60 EKMTPVSEAFILYNSTLEIVREKATREEHVGHQDQIYAKQYR---FESITELED- 115
Qy 124 FKGTPLSLYMLFNTSELREAVPEVPLVSRBELRLRLKL--KVPOHVELYQKYSNDGW-- 179
Db 116 -----HEFKFKFNASHVRENKVGMSLLHHAELRYKKQTKNMDQRMELFWKYQENGTT 170
Qy 180 -RYSNRLAPSDSPWLSFEDVTGVVQWLTRREAIEGERLSAH---SSSDSKDNTLIVE 235
Db 171 SRYLESKYIPTVDDEWMSFDTVTNENLKRABENQFGLQPACKPTPOAKD-----ID 226
Qy 236 INGNSSRRGDLATIGHM--NRPELLLMATPLERAQHLHSSRRRALDTNSYPYDVPDYA 293
Db 227 IEGF-PALRGDLASLSKENTKPYLMITSMPAERIDVTSSKKR----- 270
Qy 294 SLALDTNYCFSSSTKNCVQYLYIDFRKDLGKWKWIHEPKGYHANFCGLGPCDYIWSLDTQY 353
Db 271 --GVGOEYCFGNNGPNCVKPLINFRKDLGKWKWIHEPKGYEANYCLGNCPCYIWSMDTQY 328
Qy 354 SKVLALYNHNPASAPCCVPQALEPLPIYVYVGRPKPVQOLSNMIVRSCKCS 407
Db 329 SKVLSLYNQNNPGASISCCVPOVLEPLPIIYVYVGRKTAKVQLSNMIVRSNCSS 382

RESULT 13
TGF3_CHICK
ID TGF3_CHICK
AC P16047;
```







```
DR InterPro: IPR001839; TGFb.
DR InterPro: IPR001111; TGFb_N.
DR InterPro: IPR003911; TGF-TCFb.
DR Pfam: PF00019; TGF-beta; 1.
DR Pfam: PF00688; TGFb.propeptide; 1.
DR PRINTS: PR01423; TGFbETA.
DR PRODOM: PD000357; TGFb; 1.
DR SMART: SM00204; TGFb; 1.
DR PROSITE: PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein; Signal.
FT SIGNAL 1 22
FT PROPEP 23 264
FT CHAIN 265 376
FT DISULFID 272 280
FT DISULFID 308 373
FT DISULFID 312 375
FT DISULFID 341 341
FT CARBOHYD 76 76
FT CARBOHYD 125 125
FT CARBOHYD 167 167
FT SITE 230 232
FT SEQUENCE 376 AA; 43329 MW; 7E7FC4DA58B69681 CRC64;

Query Match 36.1%; Score 780; DB 13; Length 376;
Best Local Similarity 43.6%; Pred. No. 2.6e-60;
Matches 179; Conservative 64; Mismatches 122; Indels 46; Gaps 15;

QY 6 LRLPLLLPILMLVLTGPRPAAGLSTCKTIDMELVKKRIEIRGQILSKRLASPPSQ 65
DB 1 MRVSELLALQCLGFV--HYSGLSTCSPDLDELKKRIEIRGQILSKRLSKEPEV 58

QY 66 GDVPGP-LPEAVIALYNST---RDRVAGESVEPEPE-EDYYAKEVTR--VLKVEG 118
DB 59 DEERSONIPAEILSYVNSTVELNEQAPPEQKEDPVEEYAKEVHKFTIKLMEKNP 118

QY 119 QIYDFKGTGHSYLMFNTSELREAVPEVLLSRAELRLRLK--VEOHVELYQKYSN 176
DB 119 ---DKF-----LWFNIIDISQTLNRIISQVELRLLLITTPDGSQERLELVQVGN 167

QY 177 DSWYLSNRLIAPSDPEWLSFDVTGVVROWLTRREAEIGFRLSAHSSDSKDNLTLYEI 236
DB 168 KS-RYLSRFI--PNQKWLISFDVTQTLKDWLQSEAEQGLKWNDCDPO-KTFQLKI 223

QY 237 NGFNSGRGDLATII-GNNEPELLMATPLERAQHLHSSRRALDNTSYDPDPAASL 295
DB 224 PGLVL-VRGDTETLAVNMPPEHILVMSPLD--GNSSKSRKRQPTQDQVCTDKSD--- 277

QY 296 ALDNTYCSSTKNCVRLQYIDFRKDLGKWIHEPKGVHANFCGLGPPCIYWSLDTQYSK 355
DB 278 -----GCCVRSLYIDFRKDLGKWIHEPKGVHANFCGLGPPCIYWSLDTQYSK 324

QY 356 VLALYNQHNPGASAPCCVPOALEPLPIYVYVGRKPKVEQLSNMIVRSCKC 406
DB 325 VLALYKHHPGASAPCRVYVNLPLPIFYVYVGRKPKVEQLSNMIVKCKC 375

RESULT 10
Q9ERB7
ID Q9ERB7 PRELIMINARY; PRT; 399 AA.
AC Q9ERB7
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Transforming growth factor-beta 2 (Fragment).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP Ramesh G., Kondaliah P., Seshagiri P.B.;
RA "Differential expression and selective localization of transforming
RT growth factor-beta isoforms in the hamster uterus during estrous
cycle.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL: AY007214; AAC02247.1; -.
DR HSSP: P08112; 2TGI.
DR InterPro: IPR001839; TGFb.
DR InterPro: IPR001111; TGFb_N.
DR InterPro: IPR003911; TGF-TCFb.
DR Pfam: PF00019; TGF-beta; 1.
DR Pfam: PF00688; TGFb.propeptide; 1.
DR PRINTS: PR01423; TGFbETA.
DR PRODOM: PD000357; TGFb; 1.
DR SMART: SM00204; TGFb; 1.
DR PROSITE: PS00250; TGF_BETA_1; 1.
FT NON_TER 1 399
FT NON_TER 1 399
FT SEQUENCE 399 AA; 46078 MW; A6FF8E65EAFD5148 CRC64;

Query Match 35.2%; Score 760; DB 11; Length 399;
Best Local Similarity 42.1%; Pred. No. 1.7e-58;
Matches 177; Conservative 59; Mismatches 122; Indels 62; Gaps 15;

QY 18 LVLTPGRPAAGLSTCKTIDMELVKKRIEIRGQILSKRLASPPSGDVP-PGLPEA 76
DB 4 LLHLPV--VALSUSTCSTLMDQFMRAIRIEIRGQILSKRLKLTSPPE--DYPEDEVPE 59

QY 77 VLALYNSTRD---RVAGESVEPEPE-PEADYAKEVTRVLM---VESGNQIYDKFKGTP 128
DB 60 VISIYNSTRDQLQEKASRAACERESDEEYAKEVYKIDMPSHFSENAIPPTFY-RP 118

QY 129 HSLVMLNTSELREAVPEVLLSRAELRLRL---KLKV-EQHVLYQ-----KYSNDSW 179
DB 119 YFRIVRFVDVSMMEKNSN--LVKAEFRVFLQPKARVAEQRIELYQILKSKDLTSPTQ 175

QY 180 RYLSNRLIAPSDPEWLSFDVTGVVROWLTRREAEIGFRLSAH-----SSSDSK 228
DB 176 RYDYSVYKVRRAEGEWSLSDVDAVHEWLHKKHDKNLGFKLSLHCPCCCTFFPNNIIPNK 235

QY 229 DNTLFE---INGNSGRGDLATIHGMNR-----PFLIMATPLERAQHLHSSRRRA 279
DB 236 SEELARFAGIDGTQSHSSGHOETIKSTRKNCKTPLLMLLPSYRLESQOSNRKKR 295

QY 280 LDTNSYPDVPDPAASLADNTYCSSTKNCVRLQYIDFRKDLGKWIHEPKGVHANFC 339
DB 296 -----ALDAAVCFRNVDNCCRLPLYIDFRKDLGKWIHEPKGVHANFC 339

QY 340 LGPCPYIWSLDTQYSKVLALYNQHNPGASAPCCVPOALEPLPIYVYVGRKPKVEQLSNM 399
DB 340 AGACPYLWSSDTQHTKVLISYNTINPEASAPCCVSHDLEPLTILYIGNTPKIEQLSNM 399

RESULT 11
Q99K17
ID Q99K17 PRELIMINARY; PRT; 362 AA.
AC Q99K17
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to transforming growth factor, beta 3 (Fragment).
GN TGFb3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL: BC005513; AAH05513.1; -.
DR HSSP: P10600; 1TGJ.
DR MGD; MGI:98727; Tgfb3.
```

```

RX MEDLINE-98234044; PubMed-95731001;
RY Melby P.C., Tryon V.V., Chandrasekar B., Freeman G.L.;
RT "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and
RT analysis of cytokine mRNA expression in experimental visceral
RT leishmaniasis."
RL Infect Immun. 66:2135-2142(1998).
CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF
CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1
CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND
CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: HOMODIMER. DISULFIDE-LINKED.
CC -!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
DR EMBL: X60296; CAA42838.1; -.
DR EMBL: AF046214; AAC40099.1; -.
DR HSSP: P01137; IKLA.
DR InterPro: IPR001839; TGFb.
DR Pfam: PF00019; TGF-beta; 1.
DR ProDom: PD000357; TGFb; 1.
DR SMART: SM00204; TGFb; 1.
DR PROSITE: PS00250; TGF-BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein.
FT NON_TER 1 1
FT PROPEP <1 18
FT CHAIN 19 130
FT TRANSFORMING GROWTH FACTOR BETA 1.
FT DISULFID 25 34
FT BY SIMILARITY.
FT DISULFID 33 96
FT BY SIMILARITY.
FT DISULFID 66 129
FT BY SIMILARITY.
FT DISULFID 95 95
FT INTERCHAIN (BY SIMILARITY).
FT CONFLICT 93 93
FT G -> S (IN REF. 2).
SQ SEQUENCE 130 AA; 14997 MW; 8B41DD6CF39CCA77 CRC64;
Query Match 32.0%; Score 691.5; DB 11; Length 130;
Best Local Similarity 86.4%; Pred. No. 3.6e-53;
Matches 127; Conservative 0; Mismatches 3; Indels 17; Gaps 1;
QY 261 MATPLERAQHLHSSRRRALDTSNPYDPYASLALDINYCFSSTEKNCVRLYIDFR 320
Db 1 MATPLERAQHLOSSRHR- - - - -ALDTSNYCFSSTEKNCVRLYIDFR 43
QY 321 KDLGKWKIHEPKGYHANFCLGCPYIWSLDITQYSKVLALYNQHPGASAPCCVQALEP 380
Db 44 KDLGKWKIHEPKGYHANFCLGCPYIWSLDITQYSKVLALYNQHPGASAPCCVQALEP 103
QY 381 LPVIVYVGRKPKVEQLSNMIVRSCKS 407
Db 104 LPVIVYVGRKPKVEQLSNMIVRSYKCS 130
RESULT 13
Q95N80
ID Q95N80 PRELIMINARY; PRT; 124 AA.
AC Q95N80;
DI 01-DPC-2001 (TReMBLrel. 19, Created)
DI 01-DRC-2001 (TReMBLrel. 19, Last sequence update)
DI 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Transforming growth factor beta 1 (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Fonfara S., Groene A., Baumgaertner W.;
RT "Sequence of canine transforming growth factor beta 1 mRNA in DH82-
RT cells."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL: AF349538; AAK54072.1; -.
DR InterPro: IPR001839; TGFb.
DR Pfam: PF00019; TGF-beta; 1.
DR ProDom: PD000357; TGFb; 1.

```

```

DR SMART: SM00204; TGFβ: 1.
DR PROSITE: PS00250; TGF_BETA_1: 1.
FT NON_TER 1
FT NON_TER 124 124
SQ SEQUENCE 124 AA: 14329 MW: 21D185218E556DB CRC64;

Query Match 31.2%; Score 674.5; DB 6; Length 124;
Best Local Similarity 87.2%; Pred. No. 1.1e-51;
Matches 123; Conservative 0; Mismatches 1; Indels 17; Gaps 1;

QY 264 PLERAQHLSSRRHRALDNTNSYPDYVDYASALDNTNYCFSTSEKNCVCVQLYIDFRKDL 323
DB 1 PLERAQHLSSRROR-----ALDNTNYCFSTSEKNCVCVQLYIDFRKDL 43

QY 324 GWKWIHEPGYHANFCGCPYIWSLDTOYSKVLALYNOHNPASAPCCVQALEPLPI 383
DB 44 GWKWIHEPGYHANFCGCPYIWSLDTOYSKVLALYNOHNPASAPCCVQALEPLPI 103

QY 384 VYVGRKPKVEQLSNMIVRSC 404
DB 104 VYVGRKPKVEQLSNMIVRSC 124

RESULT 14
Q98854 PRELIMINARY; PRT; 361 AA.
AC Q98854;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Transforming growth factor beta 2 (TGF-beta 2) (Fragment).
GN TGFβ2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97354301; PubMed=9210595;
RX Sumathy K., Desai K.V., Kondiah P.;
RA 'Isolation of transforming growth factor-beta2 cDNA from a fish,
RT Cyprinus carpio by RT-PCR.';
RL Gene 191:103-107(1997).
CC -1- FUNCTION: TGF-BETA 2 HAS SUPPRESSIVE EFFECTS ON INTERLEUKIN-2
CC DEPENDENT T-CELL GROWTH.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL: U66874; AAB62983.1; -.
DR HSSP: P08112; 2TGI.
DR InterPro: IPR001839; TGFβ.
DR InterPro: IPR001111; TGFβ N.
DR InterPro: IPR003911; TGF-TGFβ.
DR Pfam: PF00019; TGF-beta; 1.
DR Pfam: PF00688; TGFβ-propeptide; 1.
DR PRINTS: PR01423; TGFβ.
DR PRODOM: PD000357; TGFβ.
DR SMART: SM00204; TGFβ: 1.
DR PROSITE: PS00250; TGF_BETA_1: 1.
KW Growth factor; Mitogen; Glycoprotein.
FT NON_TER 1
FT PROPEP <1 257
FT CHAIN 258 361
FT DISULFID 264 273
FT DISULFID 272 335
FT DISULFID 334 334
FT CARBOHYD 30 30
FT CARBOHYD 98 98
FT CARBOHYD 199 199
FT NON_TER 361 361
SQ SEQUENCE 361 AA: 41931 MW: 94D930FA970A3FD3 CRC64;

```

```

Query Match 31.2%; Score 673.5; DB 13; Length 361;
Best Local Similarity 40.4%; Pred. No. 6.1e-51;
Matches 155; Conservative 54; Mismatches 116; Indels 59; Gaps 14;

QY 52 QILSKRLASPPSQGDVPPGP--LPEAVLALYNSTRDVAGESVEPEPEPE-----ADYY 104
DB 1 QILCKLKLSCPP--EIVPEPEVSRIIAIYNSTRDLQEKANERAAATCERQRTGEYY 57

QY 105 AKETRVLM---VSENGQIYDKFGTPTSHLSYMLNTSELREAVPEPVLSSRAELRLRL- 160
DB 58 AKEVHKIDMQPFAENVIPTKHV-NPYFRRLREDVSSMEKNASN---LVKAELRIERLQ 113

QY 161 --KLKV--EORVELYQ-----KYSNDSMYRLSNRLAPSDSPMLSPFDVTGVVROWLTRRE 212
DB 114 NPKARVSEQRLELQILGHKDLTPTQRYIDSKVVRTTEGEMLSPFDVTAVSEWLHRD 173

QY 213 AIEGFRLSAHSS-----DSKDNLTLHVEINGFNSG--RRGLATI---HGMNR 255
DB 174 RNNGFKISLHCPCTFPVPSNNYIIPNKSEELANFAGIDDSFVHCGDLKMFKKRRHSQS 233

QY 256 PFLLLMATPLERAQHLSSRRHRALDNTNSYPDYVDYASALDNTNYCFSTSEKNCVCVQL 315
DB 234 PHLLMLLPYSYRLESQHKSHRQ-----KRALDAAFCFRNVQDNCCLRLSL 277

QY 316 YIDFRKDLGWKWIHEPGYHANFCGCPYIWSLDTOYSKVLALYNOHNPASAPCCVCP 375
DB 278 YIDFRKDLGWKWIHEPGYHANFCAGACPYLWSADTQHSNLTGLYNTINPEASAPCCVS 337

QY 376 QALEPLPIVYVGRKPKVEQLSNM 399
DB 338 QDLEPLTILYVIGTKPKIEQLSNM 361

RESULT 15
O02730 PRELIMINARY; PRT; 112 AA.
AC O02730; O97501;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).
GN TGFβ1 OR TGF-BETA-1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Taylor T.K., James E.R., McGonigle S., Yoho E.R.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 2-99 FROM N.A.
RA Inoue K., Kawabe Y., Kodama T.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF
CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1
CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND
CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL: AF000133; AAB53806.1; -.
DR EMBL: AB020217; BAA36950.1; -.
DR HSSP: P01137; 1KLA.
DR InterPro: IPR002400; GF_cysknot.
DR InterPro: IPR001839; TGFβ.
DR Pfam: PF00019; TGF-beta; 1.
DR PRINTS: PR00438; GFCYSKNOT.
DR PRODOM: PD000357; TGFβ.
DR SMART: SM00204; TGFβ: 1.
DR PROSITE: PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein.

```



```
FT NON_TER      1      1
FT CHAIN         1 112 TRANSFORMING GROWTH FACTOR BETA 1.
FT DISULFID      7 16  BY SIMILARITY.
FT DISULFID     15 78  BY SIMILARITY.
FT DISULFID     44 109 BY SIMILARITY.
FT DISULFID     48 111 BY SIMILARITY.
FT DISULFID     77 77  INTERCHAIN (BY SIMILARITY).
FT CONFLICT      2 3   LD -> FS (IN REF. 2).
FT CONFLICT     85 92  PLPIVYV -> ATAHRTTL (IN REF. 2).
SQ SEQUENCE    112 AA; 12795 MW; 53C5B7D4635A6F3 CRC64;

Query Match      29.6%; Score 638; DB 6; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      296 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGPCPYIWSLDTQYSK 355
Db      1 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGPCPYIWSLDTQYSK 60

QY      356 VLALYNQHNPGASAPCCVPQALEPLPIVYVYGRKPKVEQLSNMIVRSCKS 407
Db      61 VLALYNQHNPGASAPCCVPQALEPLPIVYVYGRKPKVEQLSNMIVRSCKS 112
```

Search completed: October 7, 2003, 18:04:16  
Job time : 102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 10:47:27 ; Search time 5308.15 Seconds  
(without alignments)  
10489.161 Million cell updates/sec

Title: US-10-017-372E-38  
Perfect score: 1361  
Sequence: 1 tgggtaccgagatggcgccctt.....cgattaaagcggccgcgact 1361

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

```

28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	1224.4	90.0	3206	4	PIGTGFB1A	M23703 Sus scrofa
2	1208.6	88.8	1326	6	AX338213	AX338213 Sequence
3	1205	88.5	1750	4	GGTGFB1	X12373 Porcine mRNA
4	1134.8	83.4	1605	4	SSTGFB1	Y00111 Porcine mRNA
5	1113.4	81.8	2221	4	AF461808	AF461808 Sus scrofa
6	993.8	73.0	1369	4	DOGTGFB1A	L34956 Canine tran
7	982.8	72.2	2527	6	E00973	E00973 cDNA encodi
8	982	72.2	2537	6	A06669	A06669 Synthetic m
9	976	71.7	1173	4	OATGFB1	X76916 O.aries mRNA
10	975.6	71.7	1780	9	BC000125	BC000125 Homo sapi
11	975.6	71.7	1780	9	BC001180	BC001180 Homo sapi
12	974.2	71.6	1561	9	AGMTGFB	M16658 Simian tran
13	971.4	71.4	1821	6	E03028	E03028 DNA encodin
14	971	71.3	1746	9	BC022242	BC022242 Homo sapi
15	969.8	71.3	1560	6	I06216	I06216 Sequence 2
16	969	71.2	2745	9	HSTGFB1	X02812 Human mRNA
17	967.8	71.1	1560	6	I08268	I08268 Sequence 2
18	950.4	69.8	1569	6	I06221	I06221 Sequence 3
19	945.4	69.5	1561	6	I08275	I08275 Sequence 3
20	930.8	68.4	1173	9	BT007245	BT007245 Homo sapi
21	930.8	68.4	1173	12	BT007866	BT007866 Synthetic
22	921.6	67.7	1173	4	ECRGFB1	X99438 E.caballus
23	921.4	67.7	1176	6	AX615127	AX615127 Sequence
24	919.8	67.6	1176	6	AX481432	AX481432 Sequence
25	919.8	67.6	1176	6	AX615128	AX615128 Sequence
26	915.2	67.2	1187	4	AF175709	AF175709 Equus cab
27	892.4	65.6	1597	10	AF191297	AF191297 Cavia por
28	864	63.5	1641	10	BC013738	BC013738 Mus muscu
29	850.2	62.5	1579	10	MUSTGFRNA	M13177 Mouse trans
30	850.2	62.5	2094	10	MMU009862	AJ009862 Mus muscu
31	830.6	61.0	1585	10	RNTGFB1	X52498 Rat mRNA fo
32	818.8	60.2	1117	4	BOVTGFB	M36271 Bovine tran
33	794.2	58.4	1125	10	AF480858	AF480858 Sigmodon

34	685.4	50.4	1675	6	I03310	I03310 Sequence 1
35	657.2	48.3	1376	6	AX528533	AX528533 Sequence
36	655.6	48.2	1389	6	AX528619	AX528619 Sequence
37	594.6	43.7	1352	6	AX528535	AX528535 Sequence
38	589.8	43.3	1350	6	AX528615	AX528615 Sequence
39	546.8	40.2	699	6	I05434	I05434 Sequence 4
40	357.6	26.3	650	6	AX336646	AX336646 Sequence
41	357.6	26.3	650	9	HUMTGFB4	M38449 Homo sapien
42	357.6	26.3	862	6	I03312	I03312 Sequence 3
43	354.4	26.0	489	6	AX455100	AX455100 Sequence
44	327.2	24.0	469	10	MATGFB1	X60296 M.auratus m
45	321.4	23.6	1256	5	CHKTGFB4	M31160 Gallus gall

# ALIGNMENTS

## RESULT 1

### PIGTGFB1A

LOCUS PIGTGFB1A 3206 bp mRNA linear MAM 31-MAR-1995

DEFINITION Sus scrofa transforming growth factor beta-1 mRNA, complete cds.

ACCESSION M23703

VERSION M23703.1 GI:755044

KEYWORDS transforming growth factor-beta-1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 3206)

AUTHORS Kondaiah,P., Van Obberghen-Schilling,E., Ludwig,R.L., Dhar,R., Sporn,M.B. and Roberts,A.B.

TITLE cDNA cloning of porcine transforming growth factor-beta 1 mRNAs. Evidence for alternate splicing and polyadenylation

JOURNAL J. Biol. Chem. 263 (34), 18313-18317 (1988)

MEDLINE 89054010

PUBMED 2461367

COMMENT On Apr 1, 1995 this sequence version replaced gi:341017.

Original source text: Sus scrofa (strain miniature swine) cDNA to mRNA.

## FEATURES

source

Location/Qualifiers

1. .3206

/organism="Sus scrofa"

/mol\_type="mRNA"

/strain="miniature swine"

/db\_xref="taxon:9823"

/cell\_type="peripheral blood lymphocyte"

gene

1. .3206

/gene="TGF-beta-1"

CDS

906. .2078

/gene="TGF-beta-1"

/codon\_start=1

/product="transforming growth factor-beta-1"

/protein\_id="AAA64616.1"

/db\_xref="GI:755045"

/translation="MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK

RIEAIKQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD

YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL

		KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFR
		LSAHCSCDSKDN TLHVEINGFNSGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRH
		RRALDTNYCFSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
		QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVVG RKPKVEQLSNMIVRSCKCS"
polyA_signal	3186..3191	
	/gene="TGF-beta-1"	
polyA_site	3206	
	/gene="TGF-beta-1"	
BASE COUNT	645 a 1041 c 924 g 596 t	
ORIGIN		

polyA site

BASE COUNT

## ORIGIN

Query Match 90.0%; Score 1224.4; DB 4; Length 3206;  
Best Local Similarity 95.4%; Pred. No. 2.5e-212;  
Matches 1292; Conservative 0; Mismatches 11; Indels 51; Gaps 1;

## Best Local Similarity

Matches 1292; Conservative 0; Mismatches 11; Indels 51; Gaps 1;

[illegible]

Db

Qy

Db

Qy

Dh

Qy

Dk

Qy

Dlc

Qy

Dk

Qy

Dk

Qy

Dk

Qy

Dk

Qy

Dk

Qy

```

Db      1501 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC 1560
Qy      666 TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 725
        |||
Db      1561 TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 1620
Qy      726 TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC 785
        |||
Db      1621 TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC 1680
Qy      786 TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 845
        |||
Db      1681 TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGA- 1739
Qy      846 CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA 905
        |||
Db      1740 -----GCCCTGGATA 1749
Qy      906 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 965
        |||
Db      1750 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 1809
Qy      966 TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT 1025
        |||
Db      1810 TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT 1869
Qy      1026 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1085
        |||
Db      1870 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1929
Qy      1086 TGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1145
        |||
Db      1930 TGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1989
Qy      1146 AGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205
        |||
Db      1990 AGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 2049
Qy      1206 TGATCGTGCGTTCTGTCAAGTGCAGCTGAGGCCCCGCCCCGCCCACAGCCCCGCCACCC 1265
        |||
Db      2050 TGATCGTGCGTTCTGTCAAGTGCAGCTGAGGCCCCGCCCCGCCCACAGCCCCGCCACCC 2109
Qy      1266 GGCAGGCCCGGCCCCACCCCCGCCCCGCTCACCGGGGCTGTATTTAAGGACATCGTGCCC 1325
        |||
Db      2110 GGCAGGCCCGGCCCCACCCCCGCCCCGCTCACCGGGGCTGTATTTAAGGACATCGTGCCC 2169
Qy      1326 CAAGCCCACTTGGGATCGATTAAAGCGGCCGCGA 1359
        |||
Db      2170 CAAGCCCACTTGGGATCGATTAAAGGTGGAGAGA 2203

```

# RESULT 2

AX338213

LOCUS AX338213 1326 bp DNA linear PAT 09-JAN-2002  
 DEFINITION Sequence 1 from Patent WO0181404.  
 ACCESSION AX338213  
 VERSION AX338213.1 GI:18128750

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1

AUTHORS Strober,W., Nakamura,K., Kitani,A. and Fuss,I.J.

TITLE Inducible plasmid vector encoding tgf\_g(b) and uses thereof

JOURNAL Patent: WO 0181404-A 1 01-NOV-2001;  
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

FEATURES Location/Qualifiers

source 1. .1326  
/organism="Sus scrofa"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9823"

CDS 16. .1188  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD20538.1"  
/db\_xref="GI:18128751"  
/translation="MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK  
RIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD  
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL  
KLKVEQHVELYQKYSNDSWRYL SNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR  
LSAHSSSDSKDNTLHVEINGFNSGRRGDLATI HGMNRP FLLL MATPLERAQHLHSSRH  
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT  
QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"

BASE COUNT 263 a 438 c 392 g 233 t

ORIGIN

Query Match 88.8%; Score 1208.6; DB 6; Length 1326;  
Best Local Similarity 95.4%; Pred. No. 2e-209;  
Matches 1297; Conservative 0; Mismatches 9; Indels 53; Gaps 3;

Qy	1	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	60
Db	6	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	65
Qy	61	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGA	120
Db	66	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGA	125
Qy	121	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGCCAGATTCTGTCCAAGCT	180
Db	126	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGCCAGATTCTGTCCAAGCT	185
Qy	181	TCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGT	240
Db	186	TCGGCTCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCCGT	245
Qy	241	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCC	300
Db	246	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCC	305
Qy	301	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGG	360
Db	306	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGG	365

Qy	361	CAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	420
Db	366	CAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	425
Qy	421	GTCGGAGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	480
Db	426	GTCGGAGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	485
Qy	481	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGA	540
Db	486	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGA	545
Qy	541	TTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTC	600
Db	546	TTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTC	605
Qy	601	CTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	660
Db	606	CTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	665
Qy	661	TCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAA	720
Db	666	TCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAA	725
Qy	721	CGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	780
Db	726	CGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	785
Qy	781	CCTGCTCCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCG	840
Db	786	CCTGCTCCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCG	845
Qy	841	CCGAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCT	900
Db	846	CCGA-----GCCCT	854
Qy	901	GGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACAT	960
Db	855	GGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACAT	914
Qy	961	TGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAA	1020
Db	915	TGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAA	974
Qy	1021	TTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCT	1080
Db	975	TTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCT	1034
Qy	1081	GGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGC	1140
Db	1035	GGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGC	1094
Qy	1141	GCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTC	1200
Db	1095	GCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTC	1154



Qy 1201 CAACATGATCGTGC GTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCC 1260  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1155 CAACATGATCGTGC GTTCCTGCAAGTGCAGCTGA-GCCCCGCCCCGCCACAGCCCCGCC 1213  
 Qy 1261 CACCCGGCAGGCCCCGGCCCCACCCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCG 1320  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1214 CACCCGGCAGGCCCCGGCCCCACCCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCG 1273  
 Qy 1321 TGCCCCAAGCCCACCTTGGGATCGATTAAAGCGGCCCGCA 1359  
 ||||||||||||||||||||||||||||| | |  
 Db 1274 TGCCCCAAGCCCAC-TGGGATCGATTAAAGGTGGAGAGA 1311

# RESULT 3

GGTGFB1

LOCUS GGTGFB1 1750 bp mRNA linear MAM 27-MAR-1996

DEFINITION Porcine mRNA for transforming growth factor-beta 1.

ACCESSION X12373

VERSION X12373.1 GI:63808

KEYWORDS transforming growth factor-beta 1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1750)

AUTHORS Jakowlew,S.B., Dillard,P.J., Sporn,M.B. and Roberts,A.B.

TITLE Nucleotide sequence of chicken transforming growth factor-beta 1  
 (TGF-beta 1)

JOURNAL Nucleic Acids Res. 16 (17), 8730 (1988)

MEDLINE 88335639

PUBMED 3166520

REFERENCE 2 (bases 1 to 1750)

AUTHORS Jakowlew,S.B.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-1988) Jakowlew S.B., National Institute of  
 health, National Cancer Institute, Laboratory of Chemoprevention,  
 Building 41, Room B902, Bethesda, Maryland 20892, USA

COMMENT The submitters believe that the chicken cDNA library was  
 contaminated with porcine cDNA, and that the sequence is infact  
 porcine TGF-beta-1. 27-MAR-1996.

FEATURES Location/Qualifiers

source 1. .1750  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /strain="white leghorn"  
 /db\_xref="taxon:9823"  
 /clone="pTGFB-ChX119"  
 /cell type="chondrocyte"

gene 1. .1750  
 /gene="TGF-beta 1"

5'UTR 1. .446  
 /gene="TGF-beta 1"

CDS 447. .1622  
 /gene="TGF-beta 1"  
 /codon\_start=1  
 /product="transforming growth factor"  
 /protein\_id="CAA30933.1"

```

/db_xref="GI:63809"
/db_xref="SWISS-PROT:P07200"
/translation="MPPSGPGLLPLLLPLLLVLTTPGRPAAGLSTCKTIDMELVKRK
RIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRL
KLKVEQHVELYQKYSNDSWGYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTHLVEINAGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSR
HRRALDTNYCFSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD
TQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide 1285. .1619
              /gene="TGF-beta 1"
3'UTR        1620. .1750
              /gene="TGF-beta 1"
BASE COUNT   325 a    627 c    479 g    319 t
ORIGIN

```

```

Query Match      88.5%;  Score 1205;  DB 4;  Length 1750;
Best Local Similarity  94.9%;  Pred. No. 8.8e-209;
Matches 1288;  Conservative  0;  Mismatches  15;  Indels  54;  Gaps  2;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGC 65
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      442 CCCCCATGCCGCCTTCGGGGCCTGGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGC 501

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      502 TAGTGCTGACGCCTGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG 561

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      562 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 621

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      622 TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 681

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      682 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 741

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 365
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      742 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 801

Qy      366 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      802 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 861

Qy      426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      862 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 921

Qy      486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 545
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      922 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 981

Qy      546 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 605

```

Db	982	GGGGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	1041
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1042	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	1101
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAAC--G	722
Db	1102	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGCAG	1161
Qy	723	GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	782
Db	1162	GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	1221
Qy	783	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCC	842
Db	1222	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCC	1281
Qy	843	GAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGG	902
Db	1282	GA-----GCCCTGG	1290
Qy	903	ATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTG	962
Db	1291	ATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTG	1350
Qy	963	ACTTCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATT	1022
Db	1351	ACTTCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATT	1410
Qy	1023	TCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGG	1082
Db	1411	TCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGG	1470
Qy	1083	CTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGC	1142
Db	1471	CTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGC	1530
Qy	1143	TGGAGCCACTGCCCATCGTGTAACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCA	1202
Db	1531	TGGAGCCACTGCCCATCGTGTAACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCA	1590
Qy	1203	ACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCCA	1262
Db	1591	ACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCCA	1650
Qy	1263	CCCGGCAGGCCCCGGCCCCACCCCCGCCCCCTACCGGGGCTGTATTTAAGGACATCGTG	1322
Db	1651	CCCGGCAGGCCCCGGCCCCACCCCCGCCCCCTACCGGGGCTGTATTTAAGGACATCGTG	1710
Qy	1323	CCCCAAGCCCACTTGGGATCGATTAAAGCGGCCGCGA	1359
Db	1711	CCCCAAGCCCACTTGGGATCGATTAAAGGTGGAGAGA	1747

SSTGFBR  
LOCUS SSTGFBR 1605 bp mRNA linear MAM 27-MAR-1995  
DEFINITION Porcine mRNA for transforming growth factor-beta (TFG) precursor.  
ACCESSION Y00111  
VERSION Y00111.1 GI:2129  
KEYWORDS transforming growth factor-beta.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 1605)  
AUTHORS Derynck,R. and Rhee,L.  
TITLE Sequence of the porcine transforming growth factor-beta precursor  
JOURNAL Nucleic Acids Res. 15 (7), 3187 (1987)  
MEDLINE 87174844  
PUBMED 3470708  
FEATURES Location/Qualifiers  
source 1..1605  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/tissue\_type="ovary"  
/clone\_lib="lambda gt10"  
CDS 404..1576  
/note="TFG-beta precursor (AA 1-390)"  
/codon\_start=1  
/protein\_id="CAA68291.1"  
/db\_xref="GI:2130"  
/db\_xref="SWISS-PROT:P07200"  
/translation="MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK  
RIEAIRGQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD  
YYAKEVTRVLMLESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL  
KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR  
LSAHCSCDSKDNTLHVEINGFNSGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRH  
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT  
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"  
mat\_peptide 1238..1573  
/product="mature TFG-beta"  
misc\_feature 647..655  
/note="pot. N-glycosylation site"  
misc\_feature 809..817  
/note="pot. N-glycosylation site"  
BASE COUNT 298 a 572 c 437 g 298 t  
ORIGIN

Query Match 83.4%; Score 1134.8; DB 4; Length 1605;  
Best Local Similarity 95.4%; Pred. No. 4.9e-196;  
Matches 1200; Conservative 0; Mismatches 7; Indels 51; Gaps 1;

Qy 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65  
|| ||| ||||||||||||||||||||||||||||||||||||||||  
Db 399 CCCCCATGCCGCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 458  
Qy 66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 125  
||||| ||||||||||||||||||||||||||||||||||||||||  
Db 459 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 518

QY	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCTCAAGCTTCGGC	185
Db	519	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCTCAAGCTTCGGC	578
QY	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	579	TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	638
QY	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	305
Db	639	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	698
QY	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	699	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	758
QY	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	759	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	818
QY	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	819	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	878
QY	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	879	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	938
QY	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	939	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	998
QY	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	999	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	1058
QY	666	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1059	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	1118
QY	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTTGC	785
Db	1119	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTTGC	1178
QY	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1179	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	1237
QY	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1238	-----GCCCTGGATA	1247
QY	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1248	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	1307
QY	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCT	1025

```

      |||||||
Db      1308 TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTCT 1367
      |||||||
Qy      1026 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1085
      |||||||
Db      1368 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1427
      |||||||
Qy      1086 TGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1145
      |||||||
Db      1428 TGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1487
      |||||||
Qy      1146 AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205
      |||||||
Db      1488 AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1547
      |||||||
Qy      1206 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCAC 1263
      |||||||
Db      1548 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCAC 1605

```

# RESULT 5

AF461808

LOCUS AF461808 2221 bp mRNA linear MAM 03-JAN-2002

DEFINITION Sus scrofa transforming growth factor beta 1 (TGFB1) mRNA, complete cds.

ACCESSION AF461808

VERSION AF461808.1 GI:18042250

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Polymorphism in the porcine transforming growth factor beta 1 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Direct Submission

JOURNAL Submitted (20-DEC-2001) Institute of Animal Breeding Science, University of Bonn, Endenicher Allee 15, Bonn 53115, Germany

FEATURES Location/Qualifiers

source

1. .2221  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/chromosome="6"

gene

1. .2221  
/gene="TGFB1"

CDS

1. .1173  
/gene="TGFB1"  
/note="cytokine"  
/codon\_start=1  
/product="transforming growth factor beta 1"  
/protein\_id="AAL57902.1"  
/db\_xref="GI:18042251"  
/translation="MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK"



```

Db      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCTCAGT 660
Qy      671 GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 730
      |||
Db      661 GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
Qy      731 TCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 790
      |||
Db      721 TCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
Qy      791 ATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 850
      |||
Db      781 ATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGA----- 834
Qy      851 GATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAAC 910
      |||
Db      835 -----GCCCTGGATACCAAC 849
Qy      911 TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGG 970
      |||
Db      850 TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGG 909
Qy      971 AAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTG 1030
      |||
Db      910 AAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTG 969
Qy      1031 GGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGGCTCTGTAC 1090
      |||
Db      970 GGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGGCTCTGTAC 1029
Qy      1091 AACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCA 1150
      |||
Db      1030 AACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCA 1089
Qy      1151 CTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1210
      |||
Db      1090 CTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1149
Qy      1211 GTGCGTTCCTGCAAGTGCAAGTGAGGCCCCGCCCC 1245
      |||
Db      1150 GTGCGTTCCTGCAAGTGCAAGTGAGGCCCCAAGCC 1184

```

# RESULT 6

DOGTGFB1A

LOCUS DOGTGFB1A 1369 bp mRNA linear MAM 30-OCT-1994

DEFINITION Canine transforming growth factor-beta 1 (TGFB1) mRNA, complete cds.

ACCESSION L34956

VERSION L34956.1 GI:516071

KEYWORDS homologue; transforming growth factor-beta 1.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 1369)

AUTHORS Manning,A.M., Auchampach,J.A., Drong,R.F. and Slightom,J.L.





Db	233	TCTCCAGCCCCCGAGCCAGGGGAGGTGCCCGCCGTGCCGCTGCCCGAGGCCGTGCTGG	292
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	305
Db	293	CCCTCTACAACAGCACCCGCGACCGGGTGGCGGGGAGAGCGCCGAGCCCGAGCCCGAGC	352
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	353	CCGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACACCAACA	412
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	413	AAATCTATGAGAAAGTCAAGAAAAGTCCGCACAGCATATATATGCTCTTCAACACATCAG	472
Qy	426	AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	473	AGCTCCGAGAAGCAGTGCCTGAGCCCGTCTTGCTCTCCCGGGCAGAGTTGCGCCTGCTGA	532
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCCT	545
Db	533	GGCTCAAGTTAAAAGCGGAGCAGCATGTGGAGCTGTACCAGAAATATAGCAATGATTCCCT	592
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	593	GGCGCTACCTCAGCAACCGGCTGCTGGCGCCAGCGACACGCCAGAATGGCTGTCCTTTG	652
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTCGCC	665
Db	653	ATGTCACTGGAGTCGTGAGGCAGTGGCTGAGCCATGGAGGGGAAGTCGAGGGCTTTCGCC	712
Qy	666	TCAGTGCCCACTGTTCCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	713	TCAGTGCCCACTGTTCCCTGTGACAGCAAAGATAACACACTGCAAGTAGACATTAACGGGT	772
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCTCTGC	785
Db	773	TCAGTTCCAGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGACCTTCTCTGC	832
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	833	TCCTCATGGCCACCCACTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCAGCGCCG--	890
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	891	-----GGCCCTGGACA	901
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	902	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	961
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	962	TCCGCAAGGATCTGGGCTGGAAGTGGATCCATGAGCCCAAGGGTTACCACGCTAACTTCT	1021
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085



```

FT    mat_peptide      1676..2011
FT                                /product='TGF-beta'
FT    5'UTR            1..841
FT    3'UTR            2015..2537
FT    stem_loop        37..113
FT    stem_loop        2015..2100
FT    polyA_site       2514..2519.

```

```

FEATURES             Location/Qualifiers
    source            1..2527
                        /organism="Homo sapiens"
                        /mol_type="genomic RNA"
                        /db_xref="taxon:9606"

```

```

BASE COUNT      472 a      888 c      735 g      432 t
ORIGIN

```

```

Query Match          72.2%;   Score 982.8;   DB 6;   Length 2527;
Best Local Similarity 85.1%;   Pred. No. 1.8e-168;
Matches 1144;   Conservative    0;   Mismatches 147;   Indels    53;   Gaps    2;

```

```

Qy      6  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCGGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      897 TGGTGCTGACGCCTGGCCCGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCAGGCCGTGCTCG 1076

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 305
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1077 CCCTCTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1137 CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG 1196

Qy      366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy      426 AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1257 AGCTCCGAGAAGCGGTACCTGAACCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA 1316

Qy      486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 1376

Qy      546 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

LOCUS       A06669                2537 bp          mRNA          linear          PAT 29-JUL-1993
DEFINITION  Synthetic mRNA for preTGF-Betal.
ACCESSION   A06669
VERSION     A06669.1   GI:412940
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1   (bases 1 to 2537)
  AUTHORS   .
  TITLE     NUCLEIC ACID ENCODING TGF- beta 3 AND ITS USE
  JOURNAL   Patent: WO 8912101-A 4 14-DEC-1989;
FEATURES             Location/Qualifiers
     source          1..2537
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
     CDS             842..2014
                     /codon_start=1
                     /transl_table=11
                     /product="preTGF-beta1"
                     /protein_id="CAA00588.1"
                     /db_xref="GI:412941"
                     /translation="MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNNSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLRSRGGEIEGFR
LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"
BASE COUNT      473 a       893 c       739 g       432 t
ORIGIN

```

Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGCAACC	1196
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTGTC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGTC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCT	1805
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1806	GCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1865
Qy	1086	TGTACAACCAGCACAACCCGGGCGCGTCCGGCGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1866	TGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1925
Qy	1146	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205

Db	1926	AGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1985
Qy	1206	TGATCGTGCCTTCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1986	TGATCGTGCCTTCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAG	2045
Qy	1266	GGCAGGCCCGGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	2046	GCCCGGCCCAACCCGCCCCGCCCGCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA	2105
Qy	1318	TCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	2106	CCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	2137

## RESULT 9

```

OATGFB1
LOCUS      OATGFB1              1173 bp      mRNA      linear      MAM 18-APR-1995
DEFINITION O.aries mRNA for transforming growth factor-beta I.
ACCESSION  X76916
VERSION    X76916.1   GI:496648
KEYWORDS   TGF-beta 1; transforming growth factor-beta 1.
SOURCE     Ovis aries (sheep)
  ORGANISM Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE  1
  AUTHORS  Woodall,C.J., McLaren,L.J. and Watt,N.J.
  TITLE    Sequence and chromosomal localisation of the gene encoding ovine
            latent transforming growth factor-beta 1
  JOURNAL  Gene 150 (2), 371-373 (1994)
  MEDLINE  95121932
  PUBMED   7821809
REFERENCE  2 (bases 1 to 1173)
  AUTHORS  Woodall,C.
  TITLE    Direct Submission
  JOURNAL  Submitted (24-DEC-1993) C. Woodall, Univ. of Edinburgh, Dept. of
            Veterinary Pathology, Sc. of Vet. Studies, Univ. of Edinburgh,
            Edinburgh EH9 1QH, UK
FEATURES   Location/Qualifiers
  source    1. .1173
            /organism="Ovis aries"
            /mol_type="mRNA"
            /db_xref="taxon:9940"
  CDS       1. .1173
            /codon_start=1
            /product="transforming growth factor-beta 1"
            /protein_id="CAA54242.1"
            /db_xref="GI:496649"
            /db_xref="SWISS-PROT:P50414"
            /translation="MPPSGLRLLPLLLPLLWLLMLTPGRPVAGLSTCKTIDMELVKRK
            GIEAIRGQILSKLRLASPPSQGDVPPGPLPEAILALYNSTRDRVAGESAETEPEPEAD
            YYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMFNTSELREAVPEPVLLSRADVRLRL
            KLVQEHVELYQKYSNNSWRYLSNRLAPSDSPEWLSFDVTGVVRQWLTHREEIEGFR
            LSAHCSDSKDNTLQVDINGFSSGRRGDLATIHGMRNRPFLLLMATPLERAQHLHSSRH

```





```

Db      661 GCCCACTGTTCTGTGACAGTAAGGATAACACGCTTCAAGTGGACATCAACGGGTTTCAGT 720
Qy      731 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 790
      || |||||
Db      721 TCCGGCCGCCGGGGTGACCTCGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
Qy      791 ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 850
      |||||
Db      781 ATGGCCACCCCTCTGGAGAGGGCCAGCACCTGCACAGCTCCCGCCACCGCCGA----- 834
Qy      851 GATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAAC 910
      |||||
Db      835 -----GCCCTGGACACCAAC 849
Qy      911 TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGG 970
      |||||
Db      850 TACTGCTTCAGCTCCACAGAAAAGAACTGCTGTGTTCGTCAGCTCTACATTGACTTCCGG 909
Qy      971 AAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTG 1030
      |||||
Db      910 AAGGACCTGGGCTGGAAGTGGATTTCACGAACCCAAGGGCTACCACGCCAATTTCTGCCTG 969
Qy      1031 GGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTAC 1090
      |||||
Db      970 GGGCCCTGTCCCTACATCTGGAGCCTGGACACACAGTACAGCAAGGTCCTGGCCCTGTAC 1029
Qy      1091 AACCAGCACAACCCGGGCGCGTCCGGCGGCCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA 1150
      |||||
Db      1030 AACCAGCACAACCCGGGCGCATCCGGCGGCCCGTGCTGCGTGCCCTCAGGCGCTGGAACCC 1089
Qy      1151 CTGCCCATCGTGTAACGTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1210
      |||||
Db      1090 CTGCCCATCGTGTAACGTGGGGCCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATC 1149
Qy      1211 GTGCGTTCCTGCAAGTGCAGCTGA 1234
      |||||
Db      1150 GTGCGCTCCTGCAAGTGCAGCTGA 1173

```

# RESULT 10

BC000125

LOCUS BC000125 1780 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, Similar to transforming growth factor, beta 1, clone MGC:3119 IMAGE:3351664, mRNA, complete cds.

ACCESSION BC000125

VERSION BC000125.1 GI:12652748

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1780)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,



Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	502	TGGTGCTGACGCCTGGCCGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG	561
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	562	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGC	621
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	622	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCAGGCCGCTGCTCG	681
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	682	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	741
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCTGCTAATGGTGAAAGCGGCAACC	365
Db	742	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCTGCTAATGGTGAAACCCACAACG	801
Qy	366	AAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	802	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	861
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	862	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	921
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAAATACAGCAATGATTCTT	545
Db	922	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAAATACAGCAACAATTCTT	981
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTG	605
Db	982	GGCGATACTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1041
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1042	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1101
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1102	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1161
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1162	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTTCATGGCATGAACCGGCCCTTCTCTGC	1221
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1222	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1280
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1281	-----GCCCTGGACA	1290

[illegible]

RESULT 11

LOCUS BC001180 1780 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, Similar to transforming growth factor, beta 1, clone MGC:2323 IMAGE:3356605, mRNA, complete cds.

ACCESSION BC001180

VERSION BC001180.1 GI:12654682

KEYWORDS MGC.

SOURCE            Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1780)

AUTHORS      Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory



Db	562	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	621
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	622	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	681
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	305
Db	682	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	741
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	742	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG	801
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGG	425
Db	802	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	861
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	862	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	921
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	922	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	981
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCCTTTG	605
Db	982	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTTG	1041
Qy	606	ATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1042	ATGTCAACGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1101
Qy	666	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1102	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1161
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC	785
Db	1162	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTGTC	1221
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1222	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1280
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1281	-----GCCCTGGACA	1290
Qy	906	CCAATACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1291	CCAATAATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1350
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1351	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCT	1411

Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1411	GCCTCGGGCCCTGCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1470
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1471	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1530
Qy	1146	AGCCACTGCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1531	AGCCGCTGCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1590
Qy	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCGCCCCACAGCCCCGCCACCC	1265
Db	1591	TGATCGTGCGCTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCCGCCCGCCCCGGCAG	1650
Qy	1266	GGCAGGCCCGGCCCCACCCCCGCCCGCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	1651	GCCCCGCCCCACCCCCGCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA	1710
Qy	1318	TCGTGCCCCAAGCCCACTTGGGATCGATTAA	1349
Db	1711	CCCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	1742



```

/db_xref="GI:176553"
/translation="MPPSGRLRLPLLLPLLWLLVLTSPSRPAAGLSTCKTIDMELVKRK
RIETIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLRL
KLKVEQHVELYQKYSNSNSWRYLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFR
LSAHCSCDSKDNLTQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
sig_peptide      262. .1095
                  /note="transforming growth factor-beta signal peptide"
mat_peptide      1096. .1431
                  /product="transforming growth factor-beta"
BASE COUNT      301 a    547 c    446 g    267 t
ORIGIN

```

```

Query Match          71.6%;  Score 974.2;  DB 9;  Length 1561;
Best Local Similarity 85.0%;  Pred. No. 6.8e-167;
Matches 1150;  Conservative 0;  Mismatches 143;  Indels 60;  Gaps 3;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGCTGCTGTGGCTGC 65
      || ||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      257 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCGGCTGCTGCTACCGCTGCTGTGGCTAC 316

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      317 TGGTGCTGACGCCTAGCCGGCCGGCCGGCAGGACTATCCACCTGCAAGACTATCGACATGG 376

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      377 AGCTGGTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 436

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      437 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 496

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 305
      | || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      497 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAAC 556

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      557 CGGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG 616

Qy      366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      617 AAATCTATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAG 676

Qy      426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      677 AGCTCCGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCGGGCAGAGCTGCGTCTGCTGA 736

Qy      486 GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 545
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      737 GGCTCAAGTTAAAAGTGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT 796

Qy      546 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 605
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 797 GGCGATACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTTG 856

Qy 606 ATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC 665  
|||||

Db 857 ATGTCAACGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCC 916

Qy 666 TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 725  
|||

Db 917 TTAGCGCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGT 976

Qy 726 TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACAGGCATGAACCGGCCCTTCTCTGC 785  
|||

Db 977 TCACTACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCTTTCTCTGC 1036

Qy 786 TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 845  
|||

Db 1037 TTCTCATGGCCACCCCACTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA- 1095

Qy 846 CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA 905  
|||||

Db 1096 -----GCCCTGGACA 1105

Qy 906 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 965  
|||||

Db 1106 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACT 1165

Qy 966 TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT 1025  
|||||

Db 1166 TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCT 1225

Qy 1026 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1085  
|||||

Db 1226 GCCTGGGGCCCTGTCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC 1285

Qy 1086 TGTACAACCAGCACAACCCGGGCGCGTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGG 1145  
|||||

Db 1286 TGTACAACCAGCATAACCCGGGCGCCTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGG 1345

Qy 1146 AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205  
|||||

Db 1346 AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1405

Qy 1206 TGATCGTGCGTTCTTGCAAGTGAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCC 1265  
|||||

Db 1406 TGATCGTGCGTCTCTGCAATGCAGCTGAGGCCCCGCCCCGCCACCCCGCCGAG 1465

Qy 1266 GGCAGGCCCGGCCCAACCCCGCCCGCT-----CACCGGGGCTGTATTTAAGGACA 1317  
|||

Db 1466 GCCCCGCCCCGCCCAACCCCGCCCGCTGTCTTGCCCTGGGGGCTGTATTTAAGGACA 1525

Qy 1318 -TCGTGCCCCAAGCCCACTTGGGATCGATTAAA 1349  
|||||

Db 1526 CCCGTGCCCCAAGCCCACTGGGGCCCCATTA 1558



Qy	66	TAGTGTGACGCCTGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	567	TGGTGTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG	626
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	627	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	686
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	245
Db	687	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	746
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGAGCCCGAGC	305
Db	747	CCCTGTACAACAGCACCCGCGACCGGGTGCCGGGGAGAGTGCAGAACCGAGCCCGAGC	806
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	807	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAACCCACAACG	866
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCCAGCTTATATATGCTGTTCAACACGTCCG	425
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTCTGC	785
Db	1227	TCATAACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTCTGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1346	-----GCCCTGGACA	1355

[illegible]

RESULT 14

BC022242

LOCUS BC022242 1746 bp mRNA linear PRI 04-FEB-2002

DEFINITION Homo sapiens, clone MGC:22008 IMAGE:4399762, mRNA, complete cds.

ACCESSION BC022242

VERSION BC022242.1 GI:18490115

KEYWORDS MGC.

SOURCE	Homo sapiens (human)
--------	----------------------

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1746)

AUTHORS      Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cqapbs-r@mail.nih.gov](mailto:cqapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 27 Row: e Column: 21  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```

FEATURES                      Location/Qualifiers
    source                      1. .1746
                                /organism="Homo sapiens"
                                /mol_type="mRNA"
                                /db_xref="taxon:9606"
                                /clone="MGC:22008 IMAGE:4399762"
                                /tissue_type="Duodenum, adenocarcinoma"
                                /clone_lib="NIH_MGC_88"
                                /lab_host="DH10B"
                                /note="Vector: pCMV-SPORT6"
    CDS                         370. .1542
                                /codon_start=1
                                /product="Unknown (protein for MGC:22008)"
                                /protein_id="AAH22242.1"
                                /db_xref="GI:18490116"
                                /translation="MPPSGLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
                                RIEAIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
                                YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRL
                                KLVKEQHVELYQKYSNNSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR
                                LSAHCSDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
                                RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
                                QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
BASE COUNT                    376 a      612 c      472 g      286 t
ORIGIN

```

Query Match 71.3%; Score 971; DB 9; Length 1746;  
 Best Local Similarity 84.8%; Pred. No. 2.5e-166;  
 Matches 1148; Conservative 0; Mismatches 145; Indels 60; Gaps 3;

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  |||  |||  ||  |||||  |||  |||  |||||  |||||  |||||  |||
Db     365 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 424

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    425 TGGTGCTGACGCCTGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 484

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    485 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 544

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 245
      |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    545 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG 604

```

Qy	246	CTCTTTTACAACAGTACCCGCGACCCGGGTAGCCGGGGAAAGTGTCGAACCCGGAGCCCGAGC	305
Db	605	CCCTGTACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACC GGAGCCCGAGC	664
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	665	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG	724
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	725	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	784
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	785	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	844
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	845	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	904
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	905	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	964
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	965	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1024
Qy	666	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1025	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1084
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1085	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1144
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1145	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1203
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1204	-----GCCCTGGACA	1213
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1214	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1273
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1274	TCCGCAAGGACCTCGGCTGGAAGTGATCCACGAGCCCAAGGGCTACCATGCCAATTTCT	1333
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1334	GCCTCGGGCCCTGCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1393

```

Qy      1086 TGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1145
          |||
Db      1394 TGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1453

Qy      1146 AGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205
          |||
Db      1454 AGCCGCTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1513

Qy      1206 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCC 1265
          |||
Db      1514 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGTCCC GCCCGCCCCGCCCGCCCCGGCAG 1573

Qy      1266 GGCAGGCCCCGGCCCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA 1317
          |||
Db      1574 GCCCGGCCCCACCCCGCCCGCCCGCCTTGCCCATGGGGGCTGTATTTAAGGACA 1633

Qy      1318 -TCGTGCCCCAAGCCCACTGGGATCGATTAAA 1349
          |||
Db      1634 CCCGTGCCCCAAGCCCACTGGGGCCCCATTAA 1666

```

# RESULT 15

I06216

LOCUS I06216 1560 bp DNA linear PAT 02-DEC-1994

DEFINITION Sequence 2 from Patent EP 0293785.

ACCESSION I06216

VERSION I06216.1 GI:590649

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1560)

AUTHORS Purchio,A.F., Gentry,L. and Twardzik,D.

TITLE Cloning and expression of simian transforming growth factor-SS1

JOURNAL Patent: EP 0293785-A2 2 07-DEC-1988;

FEATURES Location/Qualifiers

source 1..1560  
/organism="unknown"

BASE COUNT 301 a 547 c 442 g 267 t 3 others

ORIGIN

Query Match 71.3%; Score 969.8; DB 6; Length 1560;

Best Local Similarity 84.9%; Pred. No. 4.3e-166;

Matches 1145; Conservative 0; Mismatches 143; Indels 60; Gaps 3;

```

Qy      11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70
          |||
Db      261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320

Qy      71 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130
          |||
Db      321 CTGACGCCTAGCCGGCCGGCCCGGAGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380

Qy      131 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 190
          |||
Db      381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 440

```



Qy 191 AGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTT 250  
 |||||  
 Db 441 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCCGGCCCCGCTGCCCGAGGCCGTGCTCGCCCTG 500

Qy 251 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGCCAGAG 310  
 |||||  
 Db 501 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCGGAGCCCGAACCGGAG 560

Qy 311 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 370  
 |||||  
 Db 561 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC 620

Qy 371 TATGATAAATTCAAGGGCACCCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC 430  
 |||||  
 Db 621 TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 680

Qy 431 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 490  
 |||||  
 Db 681 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 740

Qy 491 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 550  
 |||||  
 Db 741 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 800

Qy 551 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTGATGTC 610  
 |||||  
 Db 801 TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC 860

Qy 611 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT 670  
 |||||  
 Db 861 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC 920

Qy 671 GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 730  
 |||||  
 Db 921 GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT 980

Qy 731 TCTGGCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC 790  
 |||||  
 Db 981 ACCGGCCGCGGAGGTGACCTGGCCACAATTATGGCATGAACCGGCCCTTCTGCTTCTC 1040

Qy 791 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 850  
 |||||  
 Db 1041 ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGAG----- 1095

Qy 851 GATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAAC 910  
 |||||  
 Db 1096 -----CCCTGGACACCAAC 1109

Qy 911 TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGG 970  
 |||||  
 Db 1110 TACTGCTTCAGCTCCACGGAGAAGAACTGCTNCGTGCGGCAGCTGTATATTGACTTCCGC 1169

Qy 971 AAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTG 1030  
 |||||  
 Db 1170 AAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTG 1229

Qy 1031 GGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTAC 1090

Db	1230	GGGCCCTGTCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTAC	1289
Qy	1091	AACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA	1150
Db	1290	AACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA	1349
Qy	1151	CTGCCCATCGTGTA	1210
Db	1350	CTGCCCATCGTGTA	1409
Qy	1211	GTGCGTTCTTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAG	1270
Db	1410	GTGCGCTCCTGAAAATGCAGCTGAGGCCCGCCCCGCCACCCCGGCCAGGCCCG	1469
Qy	1271	GCCCCGCCCCACCCCGCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGT	1321
Db	1470	GCCCCGCCCCACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGT	1529
Qy	1322	GCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	1530	GCCCCAAGCCCACTTGGGGCCCCATTAA	1557

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

```
Run on:      October 27, 2003, 10:47:27 ; Search time 408.887 Seconds  
            (without alignments)  
            8985.201 Million cell updates/sec
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 5105512

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```

2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	1208.6	88.8		1326	24	AAD22696	Porcine transformi
2	982.8	72.2		2527	25	ABQ76674	Androgen receptor
3	982	72.2		2537	7	AAN60972	Sequence encoding
4	982	72.2		2537	11	AAQ03301	cDNA encoding huma
5	982	72.2		2537	11	AAQ02814	Sequence of pre-TG
6	982	72.2		2537	17	AAT15720	Pre-transforming g
7	980.4	72.0		2537	15	AAQ56923	Human pre-TGF-beta
8	977.2	71.8		2537	19	AAV52933	Human pre-transfor
9	974.2	71.6		1561	11	AAQ03268	Simian transformin
10	972.6	71.5		2742	22	AAI58342	Human polynucleoti
11	971.8	71.4		1559	13	AAQ20289	Sequence encoding
12	971.4	71.4		1821	12	AAQ13392	Human pro-TGF-beta
13	970.8	71.3		1560	9	AAN81084	Coding sequence of
14	970.8	71.3		1560	11	AAQ03508	Simian Transformin
15	969	71.2		2745	16	AAT05876	cDNA encoding tran
16	969	71.2		2745	22	AAH28216	Nucleotide sequenc
17	961	70.6		1303	11	AAQ09317	Monkey transformin
18	960	70.5		4105	15	AAQ55624	TGFbeta1 5'-UTR-CD
19	958	70.4		1571	11	AAQ03269	Human transforming
20	953	70.0		1569	9	AAN81085	Coding sequence of
21	949.8	69.8		1569	11	AAQ03509	Human Transforming
22	943.8	69.3		1561	11	AAQ04908	Sequence encoding
23	940.6	69.1		1561	13	AAQ29177	TGF-beta 1/beta 2

	24	921.4	67.7	1176	25	ABV75391	TGFB1 Arg25Pro pol
	25	919.8	67.6	1176	24	ABZ35738	Human TGF beta 1 p
	26	919.8	67.6	1176	24	ABX09981	Human TGFbeta1 DNA
	27	919.8	67.6	1176	24	ABV78162	Human TGF beta 1 D
	28	919.8	67.6	1176	24	ABL91703	Human polynucleoti
	29	919.8	67.6	1176	25	ABV75392	TGFB1 Arg25Pro pol
	30	899.2	66.1	1565	13	AAQ29178	TGF-beta 1. Homo
c	31	800.6	58.8	2765	22	AAI60128	Human polynucleoti
	32	779.4	57.3	2208	13	AAQ20291	Sequence encoding
	33	777.8	57.1	2206	11	AAQ03510	Human Transforming
	34	777.8	57.1	2207	11	AAQ03511	Hybrid transformin
	35	776.2	57.0	2207	11	AAQ05127	Human TGF-Betal/TG
	36	764.8	56.2	2217	10	AAN90768	Sequence of human
	37	758.6	55.7	2200	16	AAT04115	Simian-human hybri
c	38	687.8	50.5	2773	23	AAS84421	DNA encoding novel
	39	682	50.1	834	12	AAQ12192	Sequence encoding
	40	657.2	48.3	1376	24	ABK90341	DNA encoding LAP-m
	41	655.6	48.2	1389	24	ABK90344	DNA encoding LAP-h
	42	591.4	43.5	1352	24	ABK90342	DNA encoding mIFNB
	43	589.8	43.3	1350	24	ABK90343	DNA encoding huIFN
	44	525.8	38.6	875	23	AAS70979	DNA encoding novel
	45	357.6	26.3	650	24	ABK84023	Human cDNA differe

#### ALIGNMENTS

##### RESULT 1

AAD22696

ID AAD22696 standard; cDNA; 1326 BP.

XX

AC AAD22696;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-betal) cDNA.

XX

KW Porcine; transforming growth factor beta 1; TGF-betal; gene therapy;  
 KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT CDS 16..1188

FT /\*tag= a

FT /product= "Porcine TGF-betal mutant protein"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR P-PSDB; AAE13596.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel  
PT disease in humans, comprises vector containing transforming growth  
PT factor-beta under the control of inducible promoter -

XX

PS Claim 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a  
CC gene encoding a regulatory transcription factor under the control of a  
CC promoter encoding a transforming growth factor-beta (TGF-beta). The  
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2  
CC or TGF-beta3, its variants or homologues, by transfecting a cell which  
CC is part of a host suspected of having an autoimmune disease, especially  
CC inflammatory bowel disease (IBD), under conditions such that the  
CC polypeptide encoded by the nucleic acid sequence in the vector is  
CC expressed. The vector is delivered using a delivery system. The delivery  
CC of the vector results in substantial elimination of symptoms of the  
CC autoimmune disease and increased production of IL-10 by the host. The  
CC composition is useful for treating various diseases with an autoimmune  
CC component such as multiple sclerosis, rheumatoid arthritis, systemic  
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis  
CC and psoriasis, and also for assaying the expression of a gene in a cell.  
CC The vector is further useful for screening of the effect of test  
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.  
CC The present sequence is a cDNA encoding porcine TGF-beta1 mutant.

XX

SQ Sequence 1326 BP; 263 A; 438 C; 392 G; 233 T; 0 other;

Query Match 88.8%; Score 1208.6; DB 24; Length 1326;  
Best Local Similarity 95.4%; Pred. No. 4.4e-235;  
Matches 1297; Conservative 0; Mismatches 9; Indels 53; Gaps 3;

Qy	1	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	60
Db	6	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	65
Qy	61	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGA	120
Db	66	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGA	125
Qy	121	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCT	180
Db	126	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCT	185
Qy	181	TCGGCTTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGT	240
Db	186	TCGGCTCGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCCGT	245
Qy	241	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCC	300
Db	246	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCC	305

Qy	301	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAAAGCGG	360
Db	306	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAAAGCGG	365
Qy	361	CAACCAAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	420
Db	366	CAACCAAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	425
Qy	421	GTCGAGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	480
Db	426	GTCGAGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	485
Qy	481	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAAATACAGCAATGA	540
Db	486	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAAATACAGCAATGA	545
Qy	541	TTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCCGAGTGGCTGTC	600
Db	546	TTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCCGAGTGGCTGTC	605
Qy	601	CTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	660
Db	606	CTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	665
Qy	661	TCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAA	720
Db	666	TCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAA	725
Qy	721	CGGGTTCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	780
Db	726	CGGGTTCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	785
Qy	781	CCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCG	840
Db	786	CCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCG	845
Qy	841	CCGAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCT	900
Db	846	CCGA-----GCCCT	854
Qy	901	GGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACAT	960
Db	855	GGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACAT	914
Qy	961	TGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAA	1020
Db	915	TGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAA	974
Qy	1021	TTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTT	1080
Db	975	TTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTT	1034
Qy	1081	GGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGC	1140
Db	1035	GGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGC	1094

```

Qy      1141 GCTGGAGCCACTGCCCATCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTC 1200
          |||
Db      1095 GCTGGAGCCACTGCCCATCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTC 1154

Qy      1201 CAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCGCCCCACAGCCCCGCC 1260
          |||
Db      1155 CAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA-GCCCCGCCCCGCCCCACAGCCCCGCC 1213

Qy      1261 CACCCGGCAGGCCCGGCCCAACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCG 1320
          |||
Db      1214 CACCCGGCAGGCCCGGCCCAACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCG 1273

Qy      1321 TGCCCCAAGCCCACTTGGGATCGATTAAAGCGGCCGCGA 1359
          |||
Db      1274 TGCCCCAAGCCCAC-TGGGATCGATTAAAGGTGGAGAGA 1311

```

# RESULT 2

ABQ76674

ID ABQ76674 standard; DNA; 2527 BP.

XX

AC ABQ76674;

XX

DT 26-MAR-2003 (first entry)

XX

DE Androgen receptor signalling pathway-associated DNA E00973.

XX

KW Androgen receptor; transactivation; modulator; Smad3; Smad4; Akt; TGF-B;  
KW signal transduction pathway; transforming growth factor-B; phosphatase;  
KW tensin; cytostatic; antiproliferative; cellular proliferation; cancer;  
KW E00973; ds.

XX

OS Synthetic.

XX

PN WO200282081-A2.

XX

PD 17-OCT-2002.

XX

PF 05-APR-2002; 2002WO-US11086.

XX

PR 06-APR-2001; 2001US-282266P.

PR 13-MAR-2002; 2002US-365060P.

XX

PA (UYRP ) UNIV ROCHESTER.

XX

PI Chang C;

XX

DR WPI; 2003-046871/04.

XX

PT Modulating androgen receptor activity, by administering a compound that  
PT modulates receptor activity, inhibits receptor-signal transduction  
PT pathway/receptor-coactivator interaction or changes amount or receptor  
PT -

XX

PS Disclosure; Page 225-226; 302pp; English.

XX

CC This invention describes a novel method for modulating androgen receptor

CC activity or androgen receptor-mediated transactivation activity in a  
CC cell. The method involves administering a compound which causes  
CC modulation of the androgen receptors activity and the inhibition of  
CC interaction between the receptor and a protein involved in a signal  
CC transduction pathway. The compound also inhibits the interaction between  
CC the androgen receptor and a protein selected from Smad3, Smad4, Akt,  
CC transforming growth factor (TGF)-B and phosphatase and tensin homologues  
CC deleted on chromosome 10 (PTEN) or their fragments. The compounds of the  
CC invention have cytostatic and antiproliferative activity. The obtained  
CC composition is useful for treating any disease, where uncontrolled  
CC proliferation or cellular proliferation occurs such as cancer, e.g.  
CC prostate cancer. This sequence represents the androgen receptor  
CC transactivation signalling pathway modulator E00973 described in  
CC the method of the invention.

XX

SQ Sequence 2527 BP; 472 A; 888 C; 735 G; 432 T; 0 other;

Query Match 72.2%; Score 982.8; DB 25; Length 2527;  
Best Local Similarity 85.1%; Pred. No. 2e-189;  
Matches 1144; Conservative 0; Mismatches 147; Indels 53; Gaps 2;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 125
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    897 TGGTGCTGACGCCTGGCCCCGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy    246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 305
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1077 CCCTCTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC 365
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG 1196

Qy    366 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy    426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy    486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 545
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT 1376
```



Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCCGAGTGGCTGTCTCTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAATTCT	1805
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1806	GCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCACTACAGCAAGGTCCTGGCCC	1865
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGG	1145
Db	1866	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGG	1925
Qy	1146	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1926	AGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGGAGCTGTCCAACA	1985
Qy	1206	TGATCGTGCGTTCTTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1986	TGATCGTGCGTCTCTGCAAGTGCAGCTGAGGTCCCGCCCCGCC--CCGCCCGCCCGGCC	2043
Qy	1266	GGCAGGCCCGGCCCCACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCGTGCCC	1325
Db	2044	CCACCCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCC	2103
Qy	1326	CAAGCCCACTTGGGATCGATTAAA	1349
Db	2104	CAAGCCCACTTGGGGCCCCATTAA	2127

RESULT 3

AAN60972

ID AAN60972 standard; cDNA; 2537 BP.

XX

AC AAN60972;

XX

DT 31-OCT-2002 (updated)

DT 28-OCT-1991 (first entry)

XX

DE Sequence encoding preTGF-beta.

XX

KW Transforming growth factor beta; cancer; wound healing.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT misc\_structure 37..113

FT /\*tag= a

FT /note= "Sequence can form stable hairpin loops"

FT CDS 842..2014

FT /\*tag= b

FT mat\_peptide 1676..2011

FT /\*tag= c

XX

PN EP200341-A.

XX

PD 10-DEC-1986.

XX

PF 21-MAR-1986; 86EP-0302112.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA;

XX

DR WPI; 1986-326875/50.

DR P-PSDB; AAP61468.

XX

PT TGF-beta prodn. from transformed hosts - useful esp. for treating

PT wounds (J6 2/9/86).

XX

PS Disclosure; Fig 1b; 26pp; English.

XX

CC The gene product is known to stimulate cell proliferation and  
 CC inhibit anchorage-dependent growth of a variety of human cancer cell  
 CC lines, it is esp. useful in treatment of burns and the promotion of  
 CC surface and internal wound healing. TGF-beta may be expressed from a  
 CC transformed CHO cell line.

CC (Updated on 31-OCT-2002 to add missing OS field.)

XX

SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 72.2%; Score 982; DB 7; Length 2537;

Best Local Similarity 84.9%; Pred. No. 3e-189;

Matches 1148; Conservative 0; Mismatches 145; Indels 59; Gaps 2;

QY	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	65
Db	837	CCCCCATGCGCGCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC	896
QY	66	TAGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	897	TGGTGCTGACGCCTGGCCCGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG	956
QY	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	957	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
QY	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	1076
QY	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
QY	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG	1196
QY	366	AAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
QY	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
QY	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1376
QY	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGTTTATCTTTTG	1436
QY	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
QY	666	TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
QY	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTCTGC	1616
QY	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675

QY	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
QY	906	CCAAC TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAAC TATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
QY	966	TCCGGAAGGACCTGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAATTCT	1805
QY	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1806	GCCTCGGGCCCTGCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1865
QY	1086	TGTACAACCAGCACAACCCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1866	TGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1925
QY	1146	AGCCACTGCCCATCGTG TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1926	AGCCGCTGCCCATCGTG TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1985
QY	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1986	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGTCCCCGCCCCGCCCGCCCCGCCGCGCAG	2045
QY	1266	GGCAGGCCCCGCCCCACCCCCGCCCCCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	2046	GCCCGGCCCCACCCCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA	2105
QY	1318	TCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	2106	CCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	2137

# RESULT 4

AAQ03301

ID AAQ03301 standard; DNA; 2537 BP.

XX

AC AAQ03301;

XX

DT 25-MAR-2003 (updated)

DT 05-AUG-1990 (first entry)

XX

DE cDNA encoding human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).

XX

KW Transforming growth factor-beta-1 (TGF-beta-1);

KW neoplastic cell line inhibition;

KW EGF-potentiased anchorage-independent growth;

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 842..2014

```

FT          /*tag=  a
FT  mat_peptide  1676..2011
FT          /*tag=  b
FT  misc_difference  37..113
FT          /*tag=  c
FT          /note="stable hairpin loops"
FT  misc_feature  2015..2100
FT          /*tag=  d
FT          /note="G-C rich sequence
FT          and a downstream TATA-like sequence"
XX
PN  US4886747-A.
XX
PD  12-DEC-1989.
XX
PF  13-MAR-1987;   87US-0025423.
XX
PR  13-MAR-1987;   87US-0025423.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Derynck RMA,  Goeddel DV;
XX
DR  WPI; 1990-051338/07.
DR  P-PSDB; AAR05258.
XX
PT  Nucleic acid encoding transforming growth factor-beta -
PT  cloned into expression vectors for expression in eukaryotic host
PT  cells for therapeutic use
XX
PS  Disclosure;  Fig 1b; 28pp; English.
XX
CC  It was obtained by an analysis of several overlapping cDNAs and gene
CC  fragments, leading to the detn. of a continuous sequence corresp. to the
CC  TGF-beta-1 precursor mRNA. It is useful in constructing vectors that
CC  encode biologically active transforming growth factor (TGF-beta),
CC  operably linked to DNA that encodes a secretory leader (SL). It, or a
CC  nucleic acid capable of hybridising with it, can also be labelled and
CC  used in diagnostic assays for DNA or mRNA encoding TGF-beta or related
CC  proteins.
CC  (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ  Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

```

```

Query Match          72.2%;  Score 982;  DB 11;  Length 2537;
Best Local Similarity  84.9%;  Pred. No. 3e-189;
Matches 1148;  Conservative  0;  Mismatches 145;  Indels  59;  Gaps  2;

```

```

Qy          6  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
              ||   |||  |||  ||  ||||| |||||  ||||| ||||| ||||| ||||| |||||
Db          837  CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy          66  TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 125
              |  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          897  TGGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTATCCACCTGCAAGACTATCGACATGG 956

Qy          126  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185

```

Db	957	 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGCCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGCCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG	1196
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGC	1616
Qy	786	TCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025

Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCT	1805
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1806	GCCTCGGGCCCTGCCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1865
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1866	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1925
Qy	1146	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1926	AGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1985
Qy	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1986	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAG	2045
Qy	1266	GGCAGGCCCGGCCCCACCCCCGCCCCCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	2046	GCCCCGCCCCACCCCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA	2105
Qy	1318	TCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	2106	CCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	2137

FT /label=

```
FT          /*tag= e
FT  misc_feature  863..911
FT          /*tag= f
FT          /label=hydrophobic domain
```

PN WO8912101-A.

PD 14-DEC-1989.

PF 08-JUN-1988; 88WO-US01945.

PR 08-JUN-1988; 88WO-US01945.

PA (GETH ) GENENTECH INC.

PI DERNYCK RMA, GOEDDEL DV;

DR WPI; 1990-007474/01.

DR P-PSDB; AAR04034.

PT Nucleotide sequence encoding transforming growth factor beta-3 -used as a  
PT probe, or to produce TGF beta 3, for growth inhibition of certain normal  
PT and neoplastic cells, eg A549.

PS Disclosure; Fig. 1b; 61pp; English.

Sequence encodes the 390 amino acid (AA) precursor transforming growth factor-beta 1 (pre-TGF-beta 1) polypeptide. The 5' untranslated region of the TGF-beta 1 mRNA is 841 bases long, is purine rich and has a region of potential secondary structure. The TATA-like sequence in the 3' untranslated region of the gene is presumably a polyadenylation signal. Mature TGF-beta 1 comprises the C-terminal 112 AA's of pre-TGF-beta 1 and is cleaved at the Arg-Arg dipeptide preceding its NH2 terminus. The nucleic acid encoding the second subtype of TGF-beta (TGF-beta 3) is useful as a probe or to produce TGF-beta 3 for inhibition of growth of normal and neoplastic cells.

CC (Updated on 31-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

SO Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 72.2%; Score 982; DB 11; Length 2537;

Best Local Similarity 84.9%; Pred. No. 3e-189;

Matches 1148; Conservative 0; Mismatches 145; Indels 59; Gaps 2;

QY 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65

Db 837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy 66 TAGTGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG 125

Db 897 TGGTGCTGACGCCCTGGCCCGCCGGCCGCGGGACTATCCACCTGCAAGACTATCGACATGG 956

QY 126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 185

Db 957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016



Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
Qy	906	CCAATACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAATAATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCT	1805

Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTC	1085
Db	1806	GCCTCGGGCCCTGCCCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1865
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1866	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1925
Qy	1146	AGCCACTGCCCATCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1926	AGCCGCTGCCCATCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1985
Qy	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1986	TGATCGTGCGCTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGGCAG	2045
Qy	1266	GGCAGGCCCGGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	2046	GCCCGGCCCCACCCCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA	2105
Qy	1318	TCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	2106	CCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	2137

## RESULT 6

ID AAT

AC

DT

DT

DE

KW

YY

YY

ET

11

FI  
DD

ET

FT

FT

FT

FT repeat\_region 2015..2100  
 FT /\*tag= e  
 FT /note= "GC-rich region; possibly responsible for the  
 FT fact 3'UTR of mRNA could not be cloned as cDNA;  
 FT may be important for transcription efficiency"  
 FT repeat\_unit 2019..2023  
 FT /\*tag= f  
 FT TATA\_signal 2094..2100  
 FT /\*tag= g  
 FT /note= "TATA-like sequence; no evidence that this  
 FT functions a promoter"  
 FT polyA\_signal 2514..2520  
 FT /\*tag= h  
 FT misc\_signal 2529..2536  
 FT /\*tag= i  
 FT /note= "consensus sequence immediately precedes  
 FT polyA-tail (Benoist et al)"  
 XX  
 PN US5482851-A.  
 XX  
 PD 09-JAN-1996.  
 XX  
 PF 05-NOV-1993; 93US-0147364.  
 XX  
 PR 13-MAR-1987; 87US-0025423.  
 PR 22-MAR-1985; 85US-0715142.  
 PR 04-AUG-1989; 89US-0389929.  
 PR 04-MAR-1992; 92US-0845893.  
 PR 05-NOV-1993; 93US-0147364.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Derynck RMA, Goeddel DV;  
 XX  
 DR WPI; 1996-076891/08.  
 DR P-PSDB; AAR90827.  
 XX  
 PT New recombinant human transforming growth factor-beta prods. - produced  
 PT using Chinese hamster ovary cells, for use in diagnostic applications  
 PT or in therapy  
 XX  
 PS Example 3; Fig 1; 26pp; English.  
 XX  
 CC The cDNA encodes the pre-transforming growth factor (TGF) beta 1 protein.  
 CC The nucleotide sequence was obt'd. by an analysis of several overlapping  
 CC cDNAs and gene fragments. The DNA is useful for the recombinant  
 CC production of TGF beta 1, which can be used in, e.g. wound healing.  
 CC (Revised entry submitted to correct sequence analysis breakdown.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 72.2%; Score 982; DB 17; Length 2537;  
 Best Local Similarity 84.9%; Pred. No. 3e-189;  
 Matches 1148; Conservative 0; Mismatches 145; Indels 59; Gaps 2;

Qy 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65

Db	837	CCCCCATGCCGCCCTCCGGGCTGCCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC	896
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	897	TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	957	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCCGCTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCCGCTGCTAATGGTGGAACCCACAACG	1196
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGTCCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGTCCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	1376
Qy	546	GGCGTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905

```

Db      1676 -----GCCCTGGACA 1685
Qy      906 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 965
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1686 CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT 1745
Qy      966 TCCGGAAGGACCTGGGCTGGAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCT 1025
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1746 TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCT 1805
Qy      1026 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1085
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1806 GCCTCGGGCCCTGCCCCACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC 1865
Qy      1086 TGTACAACCAGCACAAACCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1145
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1866 TGTACAACCAGCATAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1925
Qy      1146 AGCCACTGCCCATCGTGTAATACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1926 AGCCGCTGCCCATCGTGTAATACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1985
Qy      1206 TGATCGTGCGTTCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCC 1265
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1986 TGATCGTGCGTCTCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCCGCCCCGCCCCGGCAG 2045
Qy      1266 GGCAGGCCCCGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA 1317
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2046 GCCCGGCCCCACCCCCGCCCCGCCCCGCTGCTTGCCCATGGGGGCTGTATTTAAGGACA 2105
Qy      1318 TCGTGCCCCAAGCCCACTTGGGATCGATTAAA 1349
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2106 CCGTGCCCCAAGCCCACTTGGGGCCCCATTAA 2137

```

# RESULT 7

AAQ56923

ID AAQ56923 standard; cDNA; 2537 BP.

XX

AC AAQ56923;

XX

DT 25-MAR-2003 (updated)

DT 09-JUL-1994 (first entry)

XX

DE Human pre-TGF-beta-1.

XX

KW TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;

KW transforming growth factor beta-3; recombinant; wound healing;

KW vulnerary; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT misc\_structure 47..113

FT /\*tag= a

FT /note= "possible hairpin loop region"

FT CDS 842..2014

```

FT          /*tag= b
FT  mat_peptide  1676..2011
FT          /*tag= c
FT  polyA_signal  2515..2521
FT          /*tag= d
XX
PN  US5284763-A.
XX
PD  08-FEB-1994.
XX
PF  04-MAR-1992;  92US-0845893.
XX
PR  22-MAR-1985;  85US-0715142.
PR  13-MAR-1987;  87US-0025423.
PR  04-AUG-1989;  89US-0389929.
PR  04-MAR-1992;  92US-0845893.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Derynk RMA,  Goeddel DV;
XX
DR  WPI; 1994-056343/07.
DR  P-PSDB; AAR46227.
XX
PT  Nucleic acid sequences encoding transforming growth factor-beta -
PT  diagnostic probes, and for use in therapeutics
XX
PS  Disclosure; Fig 1b; 25pp; English.
XX
CC  cDNA sequences were determined for human pre-TGF-beta-1 (AAQ56923),
CC  pig TGF-beta-3 (AAQ56925) and human TGF-beta-3 (AAQ56926), and the
CC  corresponding amino acid sequences were determined (AAR46227-29,
CC  respectively). A genomic fragment corresponding to a human TGF-
CC  beta-1 exon (AAQ56924) was also isolated and its amino acid sequence
CC  determined (AAR46230). The sequences have been used in the
CC  construction of vectors for the expression of recombinant TGF-
CC  beta.
CC  (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ  Sequence 2537 BP; 473 A; 890 C; 742 G; 432 T; 0 other;

Query Match          72.0%;  Score 980.4;  DB 15;  Length 2537;
Best Local Similarity  84.8%;  Pred. No. 6.2e-189;
Matches 1147;  Conservative  0;  Mismatches 146;  Indels  59;  Gaps  2;

Qy          6  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC  65
   ||  ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          837  CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC  896

Qy          66  TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG  125
   |  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          897  TGGTGCTGACGCCTGGCCCCGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG  956

Qy          126  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC  185
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          957  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC  1016

```

Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACGGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	366	AAATCTATGTATAAATTCAGGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAGAGCAGGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATCTCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCAACGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGATCCACGAGCCAAGGGCTACCATGCCAATTCT	1805
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085

Db	1806	GCCTCGGGCCCTGCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1865
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1866	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1925
Qy	1146	AGCCACTGCCCATCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1926	AGCCGCTGCCCATCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1985
Qy	1206	TGATCGTGCGTTCTTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1986	TGATCGTGCGCTCTTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAG	2045
Qy	1266	GGCAGGCCCGGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	2046	GCCCGGCCCCACCCCCGCCCCGCCCCGCTTGCCCATGGGGGCTGTATTTAAGGACA	2105
Qy	1318	TCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	2106	CCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	2137

AAV52933

XX

XX

DT 21-DEC-1998 (first entry)

XX

XX

XX

XX

FT CDS 842..2014

```
FT          /*tag=  a
```

FT mat peptide 1676..2011

```
FT                               /*tag=  b
```

FT	stem loop	37..113
----	-----------	---------

FT /\*tag= b

```
FT      /note= "putative stable hairpin loop"
```

```
FT    misc_feature    2015..2100
```

FT /\*tag= c

```
FT                               /note= "GC-rich sequence"
```

FT	polyA_signal	2514..2520
----	--------------	------------

```
FT                                /*tag= d
```

XX

XX

PD 01-SEP-1998.

XX



PF 30-MAY-1995; 95US-0454468.  
XX  
PR 13-MAR-1987; 87US-0025423.  
PR 22-MAR-1985; 85US-0715142.  
PR 04-AUG-1989; 89US-0389929.  
PR 04-MAR-1992; 92US-0845893.  
PR 05-NOV-1993; 93US-0147364.  
PR 30-MAY-1995; 95US-0454468.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1998-494840/42.

DR P-PSDB; AAW78785.

XX

PT DNA encoding transforming growth factor-beta precursor sequence -  
PT useful for analysis to perform manipulations to increase yield of  
PT recombinant production of the protein

XX

PS Example 3; Fig 1B 1-3; 26pp; English.

XX

CC This nucleotide sequence codes for the human transforming growth  
CC factor-beta 1 precursor (preTGF-beta 1, see AAW78785). It is a  
CC composite of overlapping cDNA clones isolated from different cDNA  
CC libraries (placenta, A172 glioblastoma, HT1080 fibroblastoma) using  
CC TGF-beta exon (see AAV52936) restriction fragments as probes.  
CC The 3' region of the sequence was determined using cloned genomic  
CC DNA. The invention relates to the recombinant production of  
CC TGF-beta. Biologically active TGF-beta is defined as being capable  
CC of inducing EGF-potentiased anchorage independent growth of target  
CC cell lines and/or growth inhibition of neoplastic cell lines.  
CC Nucleic acids encoding TGF-beta have been isolated and cloned into  
CC vectors which are replicated in bacteria and expressed in  
CC eukaryotic cells. TGF-beta recovered from transformed cells is  
CC used in known therapeutic applications. TGF-beta nucleic acids are  
CC also useful in diagnosis and identification of TGF-beta clones.  
CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 2537 BP; 475 A; 895 C; 736 G; 431 T; 0 other;

Query Match 71.8%; Score 977.2; DB 19; Length 2537;  
Best Local Similarity 84.7%; Pred. No. 2.8e-188;  
Matches 1145; Conservative 0; Mismatches 148; Indels 59; Gaps 2;

Qy 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65  
|| ||| |||| || ||||||||| ||||||||| ||||||||| |||  
Db 837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896  
  
Qy 66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125  
| ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 897 TGGTGCTGACGCCTGGCCCGCCGGCCCCGGGACTATCCACCTGCAAGACTATCGACATGG 956  
  
Qy 126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185  
||| ||||||||| ||||||||| ||||||||| ||||||||| ||||| |||||  
Db 957 AGCAGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCAACCGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCAACCGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGAGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTGTC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1676	-----GCCCTGGACA	1685
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1686	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1745
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTCT	1025
Db	1746	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAATTCT	1805
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085



```

XX
DR      WPI; 1990-038499/06.
DR      P-PSDB; AAR03743.
XX
PT      Inhibition of proliferation of epidermal cells -
PT      used to treat psoriasis by contacting cells with compositions
PT      containing transforming growth factor-beta.
XX
PS      Disclosure; fig 1; 20pp; English.
XX
CC      TGF-beta may be used in the treatment of hyperplasia
CC      associated with acanthosis-categorised skin diseases, and
CC      in alleviating psoriatic symptoms associated with cytokine-
CC      induced phenomena. See also AAQ03269 and AAR03750.
CC      (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ      Sequence 1561 BP; 301 A; 547 C; 446 G; 267 T; 0 other;

```

Query Match 71.6%; Score 974.2; DB 11; Length 1561;  
Best Local Similarity 85.0%; Pred. No. 1.1e-187;  
Matches 1150; Conservative 0; Mismatches 143; Indels 60; Gaps 3;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	65
Db	257	CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC	316
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGCCGGAAGTGTCCACCTGCAAGACCATCGACATGG	125
Db	317	TGGTGCTGACGCCTAGCCGGCCGGCCGCAGGACTATCCACCTGCAAGACTATCGACATGG	376
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	377	AGCTGGTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	436
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	437	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	496
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	497	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCCGAGCCCGAAC	556
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGGTGCTAATGGTGGAAAGCGGCAACC	365
Db	557	CGGAGGCCGACTACTACGCCAAGGAGGTCAACCGGTGCTAATGGTGGAAACCCACAACG	616
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	617	AAATCTATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAG	676
Qy	426	AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	677	AGCTCCGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	736
Qy	486	GGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTTCCT	545
Db	737	GGCTCAAGTTAAAAAGTGGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTTCCT	796

Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGA	605
Db	797	GGCGATACTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTTG	856
Qy	606	ATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	857	ATGTCAACGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCC	916
Qy	666	TCAGTGCCCACTGTTCCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	917	TTAGCGCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGT	976
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	785
Db	977	TCACTACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCCTGC	1036
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1037	TTCTCATGGCCACCCCACTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-	1095
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1096	-----GCCCTGGACA	1105
Qy	906	CCAATACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1106	CCAATACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACT	1165
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1166	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCT	1225
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1226	GCCTGGGGCCCTGTCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1285
Qy	1086	TGTACAACCAGCACAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1286	TGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1345
Qy	1146	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1346	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1405
Qy	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCAAGCCCCGCCACCC	1265
Db	1406	TGATCGTGCGTTCCTGCAATGCAGCTGAGGCCCCGCCCCGCCCGCCCCACCCCGGCAG	1465
Qy	1266	GGCAGGCCCGGCCCAACCCCGCCCGCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	1466	GCCCGGCCCGGCCCAACCCCGCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACA	1525
Qy	1318	-TCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1349
Db	1526	CCCGTGCCCCAAGCCCACTTGGGGCCCCATTAA	1558

RESULT 10

AAI58342

ID AAI58342 standard; cDNA; 2742 BP.

XX

AC AAI58342;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 545.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia; ss.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR P-PSDB; AAM39186.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

PS Claim 1; SEQ ID NO 545; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic







DT 25-MAR-2003 (updated)  
 DT 16-APR-1992 (first entry)  
 XX  
 DE Sequence encoding simian transforming growth factor (TGF) beta-1.  
 XX  
 KW Hypertension therapy; hypotensive agent; blood pressure modulator;  
 KW ss.  
 XX  
 OS Monkey.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 262..282  
 FT /\*tag= a  
 FT sig\_peptide 283..324  
 FT /\*tag= b  
 FT CDS 325..1098  
 FT /\*tag= c  
 FT mat\_peptide 1099..1436  
 FT /\*tag= d  
 XX  
 PN WO9119513-A.  
 XX  
 PD 26-DEC-1991.  
 XX  
 PF 20-JUN-1991; 91WO-US04449.  
 XX  
 PR 20-JUN-1990; 90US-0541221.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Oleson FB, Comereski CR;  
 XX  
 DR WPI; 1992-024199/03.  
 DR P-PSDB; AAR20124.  
 XX  
 PT Use of transforming growth factor (TGF)-beta and their  
 PT antagonists - for modulating blood pressure, for treating  
 PT hypertension and hypotension  
 XX  
 PS Disclosure; Fig 1; 42pp; English.  
 XX  
 CC A new method for treating hypertension comprises administering a  
 CC transforming growth factor (TGF)-beta to an individual at a dose  
 CC effective for lowering blood pressure; the TGF-beta may be e.g.  
 CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-  
 CC beta1 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-  
 CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2  
 CC complex.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1559 BP; 300 A; 546 C; 446 G; 267 T; 0 other;  
  
 Query Match 71.4%; Score 971.8; DB 13; Length 1559;  
 Best Local Similarity 85.1%; Pred. No. 3.2e-187;  
 Matches 1146; Conservative 0; Mismatches 142; Indels 59; Gaps 3;  
  
 Qy 11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70

Accession	Contig	Position	Sequence	Length
Db	261	320	ATGCCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	320
Qy	71	130	CTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	130
Db	321	380	CTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	380
Qy	131	190	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	190
Db	381	440	GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	191	250	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	250
Db	441	500	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	251	310	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAG	310
Db	501	560	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCGAGCCCGAACCGGAG	560
Qy	311	370	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	370
Db	561	620	GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	371	430	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	430
Db	621	680	TATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	431	490	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	490
Db	681	740	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	491	550	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAAATACAGCAATGATTCTGGCGC	550
Db	741	800	AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAAATACAGCAACAATTCTGGCGA	800
Qy	551	610	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	610
Db	801	860	TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	611	670	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCCTCAGT	670
Db	861	920	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	920
Qy	671	730	GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	730
Db	921	980	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	731	790	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	790
Db	981	1040	ACCGGCCGCCGAGGTGACCTGGCCACAATTATGGCATGAACCGGCCCTTCTGCTTCTC	1040
Qy	791	850	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGTCCCGGCACCGCCGAGCCCTG	850
Db	1041	1094	ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	851	910	GATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAAC	910

```

Db      1095 -----GCCCTGGACACCAAC 1109
Qy      911 TACTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGG 970
      |||
Db      1110 TACTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGC 1169
Qy      971 AAGGACCTGGGCTGGAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTG 1030
      |||
Db      1170 AAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTG 1229
Qy      1031 GGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGCTGTAC 1090
      |||
Db      1230 GGGCCCTGTCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGCTGTAC 1289
Qy      1091 AACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA 1150
      |||
Db      1290 AACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA 1349
Qy      1151 CTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1210
      |||
Db      1350 CTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1409
Qy      1211 GTGCGTTCCTGCAAGTGCAAGTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAG 1270
      |||
Db      1410 GTGCGTTCCTGCAAAATGCAGTGAGGCCCCGCCCCGCCACCCGGCAGGCCGG 1469
Qy      1271 GCCCGGCCACCCCGCCGC-----CTCACCGGGCTGTATTTAAGGACA-TCGTG 1322
      |||
Db      1470 CCCC GCCCACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTG 1529
Qy      1323 CCCCAAGCCCACTTGGGATCGATTAAA 1349
      |||
Db      1530 CCCCAAGCCCACCTGGGGCCCATTA 1556

```

RESULT 12

AAQ13392

ID AAQ13392 standard; DNA; 1821 BP.

XX

AC AAQ13392;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1 gene.

XX

KW Osteogenetic; tumoricidal; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 512..1684

FT /\*tag= a

FT sig\_peptide 512..598

FT /\*tag= b

FT misc\_RNA 599..1684

FT /\*tag= c

FT /note= "pro-TGF-beta 1"

```
FT      mat_peptide      1346..1684
FT      /*tag= e
FT      /note= "TGF-beta 1"
```

XX  
PN JP03180192-A.

XX  
PD 06-AUG-1991.

XX  
PF 07-DEC-1989: 89JP-0318243.

XX  
PR 07-DEC-1989: 89JP-0318243.

PA (KIRI ) KIRIN BREWERY KK.

XX  
DB WPT: 1991-271579/37

DR W11, 1991 271379/  
DR P-PSDB; AAR13813.  
XX

PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by  
PT preparing DNA chain contg. base sequence coding for human  
PT pre:pro-TGF-beta 1, forming expression vector etc.

XX  
PS Claim 1: Fig 1; 16pp; Japanese.

XX  
CC The DNA sequence encodes human prepro-TGF-beta 1 which can be  
CC produced by recombinant methods, it has osteogenetic and  
CC tumoricidal activity.

XX  
SO Sequence 1821 BP; 326 A; 679 C; 508 G; 308 T; 0 other;

Query Match 71.4%; Score 971.4; DB 12; Length 1821;  
Best Local Similarity 85.0%; Pred. No. 3.9e-187;  
Matches 1135; Conservative 0; Mismatches 141; Indels 59; Gaps 2;

Qy           6 CCGAGATGGCGCCTTCG GGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC         65  
            ||       ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |  
Db          507 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCGGCTGCTGCTACCGCTGCTGTGGCTAC         566

Qy           66 TAGTGCTGACGCCTGGCCGCCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 125  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db         567 TGGTGTGACGCCTGGCCGCCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 626

Qy 126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 627 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGC 686

Qy 186 TTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 245  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 687 TGGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCCGCTGCCCAGGCCGTGCTCG 746

Qy            246 CTCTTTACAACAGTACCCGCGACC GG GTAGCCGGGGAAAGTGTCGAACCGGAGCCC GAGC     305  
             | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db            747 CCCTGTACAACAGCACCCGCGACC GG GTGGCCGGGGAGAGTGCAGAACCGGAGCCC GAGC     806

QY            306 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAAGCGGCAACC     365  
             | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db            807 CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAACCCACAACG     866

Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGC	785
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1346	-----GCCCTGGACA	1355
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1356	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1415
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1416	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAATTCT	1475
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1476	GCCTCGGGCCCTGCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1535
Qy	1086	TGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGG	1145
Db	1536	TGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGG	1595
Qy	1146	AGCCACTGCCCATCGTGTAACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1596	AGCCGCTGCCCATCGTGTAACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1655
Qy	1206	TGATCGTGCGTTCCTGCAAGTGACAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCC	1265

Db 1656 TGATCGTGCCTCCTGCAAGTCAGCTGAGGTCCCGCCCCGCCCGCCCCGGCAG 1715

Qy 1266 GGCAGGCCCGGCCCAACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA 1317

Db 1716 GCCCGGCCCAACCCCGCCCGCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA 1775

Qy 1318 TCGTGCCCCAAGCCC 1332

Db 1776 CCGTGCCCCCAAGCC 1790

RESULT 13

AAN81084

ID AAN81084 standard; cDNA; 1560 BP.

XX

AC AAN81084;

XX

DT 25-MAR-2003 (updated)

DT 09-OCT-1990 (first entry)

XX

DE Coding sequence of simian transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; tumour treatment; ss cDNA.

XX

OS Cercopithecus aethiops.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT CDS 261..1433

FT /\*tag= a

FT sig peptide 282..323

FT /\*tag= b

FT mat peptide 1095..1433

FT /\*tag= c

XX

PN EP293785-A.

XX

PD 07-DEC-1988.

XX

PF 27-MAY-1988; 88EP-0108528.

XX

PR 29-MAY-1987; 87US-0055662.

PR 25-JAN-1988; 88US-0147842.

XX

PA (ONCO ) ONCOGEN.

PA (BRIM ) BRISTOL-MYERS CO.

XX

PI Purchio AG, Gentry L, Twardzik D;

XX

DR WPI; 1988-347488/49.

DR P-PSDB; AAP80647.

XX

PT    Prodn. of simian transforming growth factor beta-1 - by culturing

PT transfected eucaryotic cells, and new precursor proteins, useful for

PT treating tumours.

XX

PS Disclosure; Page ?; pp; English.

```

XX      The cDNA is prepd. from African green monkey cell line BSC-40 and is
CC      expressed in eukaryotic cells in plasmid pSV2. There is 100% homology
CC      between mature simian and human TGF-beta 1. The plasmid also contains
CC      the SV40 promoter and a selection marker, esp. DHFR.
CC      (Updated on 25-MAR-2003 to correct PA field.)
CC      (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Query Match          71.3%;   Score 970.8;   DB 9;   Length 1560;
Best Local Similarity 85.0%;   Pred. No. 5.1e-187;
Matches 1146;   Conservative    0; Mismatches 142; Indels    60; Gaps    3;

Qy       11 ATGGCGCCTTCGCGGCTGCGGCTCTTGCCGCTGCTGCTGCCGTGCTGTGGCTGCTAGTG 70
        ||| |||| | | ||||||||| | | ||||||||| | | ||||||||| | | |||
Db       261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGTGTGTGGCTACTGGTG 320

Qy       71 CTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130
        ||||| ||||| | | ||||||||| | | ||||||||| | | ||||||||| | | |||
Db       321 CTGACGCCTAGCCGGCCGGCCGCGGAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380

Qy       131 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 190
        ||||| ||||| | | ||||||||| | | ||||||||| | | ||||||||| | | |||
Db       381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 440

Qy       191 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 250
        ||||| ||||| | | ||||||||| | | ||||||||| | | ||||||||| | | |||
Db       441 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCGAGGCCGTGCTCGCCCTG 500

Qy       251 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAG 310
        ||||| ||||| | | ||||||||| | | ||||||||| | | ||||||||| | | |||
Db       501 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGC GGAGCCGAGCCCAACCGGAG 560

Qy       311 GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 370
        || ||||| ||||| | | ||||||||| | | ||||||||| | | ||||||||| | | |||
Db       561 GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACGAAATC 620

Qy       371 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 430
        ||||| || ||||| | | | ||||| | | ||||| | | ||||| | | |||||
Db       621 TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 680

Qy       431 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 490
        || ||||| | | || ||||| | | ||||| | | ||||| | | ||||| | | |||||
Db       681 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTC 740

Qy       491 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 550
        ||||| ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | |||||
Db       741 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA 800

Qy       551 TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC 610
        ||||| ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | |||||
Db       801 TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGAGTGGTTGTCTTTGATGTC 860

Qy       611 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT 670
        ||||| ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | |||||
Db       861 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTCGCCTTAGC 920

```





DT 09-JAN-2003 (updated)  
 DT 14-AUG-1990 (first entry)  
 XX  
 DE Simian Transforming growth factor - Beta1.  
 XX  
 KW HIV; AIDS; SIV; vaccine; AZT; CD4; cytokines; growth  
 KW factors; ds.  
 XX  
 OS Cebus apella.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 267..1437  
 FT /\*tag= a  
 FT mat\_peptide 1103..1437  
 FT /\*tag= b  
 XX  
 PN EP356935-A.  
 XX  
 PD 07-MAR-1990.  
 XX  
 PF 25-AUG-1989; 89EP-0115719.  
 XX  
 PR 25-AUG-1988; 88US-0236698.  
 XX  
 PA (ONCO ) ONCOGEN LP.  
 XX  
 PI Brankovan V, Lioubin M, Purchio A;  
 XX  
 DR WPI; 1990-068723/10.  
 DR P-PSDB; AAR05663.  
 XX  
 PT Compsns. contg. transforming growth factor beta -  
 PT used for inhibitions of HIV infection and replication in vivo.  
 XX  
 PS Disclosure; Fig 1; 20pp; English.  
 XX  
 CC TGF-beta may be used in vivo to prevent formation of syncytia and  
 CC inhibit HIV infection. TGF may also be used with other HIV treatments  
 CC (AZT, soluble CD4 etc.).  
 CC (Updated on 09-JAN-2003 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;  
  
 Query Match 71.3%; Score 970.8; DB 11; Length 1560;  
 Best Local Similarity 85.0%; Pred. No. 5.1e-187;  
 Matches 1146; Conservative 0; Mismatches 142; Indels 60; Gaps 3;  
  
 Qy 11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70  
 ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320  
  
 Qy 71 CTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 321 CTGACGCCTAGCCGGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380  
  
 Qy 131 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 190

Db	381		GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	191		AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	250
Db	441		AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	251		TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGCCAGAG	310
Db	501		TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCCGAGCCCGAACCGGAG	560
Qy	311		GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC	370
Db	561		GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACGAAATC	620
Qy	371		TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	430
Db	621		TATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	431		CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	490
Db	681		CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	491		AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	550
Db	741		AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	551		TACCTCAGCAACCGGCTGCTGGCCCCCAGTGAAGTCAACCGGAGTGGCTGTCTTTGATGTC	610
Db	801		TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	611		ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	670
Db	861		ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	920
Qy	671		GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	730
Db	921		GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGACATCAACGGGTTCACT	980
Qy	731		TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	790
Db	981		ACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCTGCTTCTC	1040
Qy	791		ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	850
Db	1041		ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGAG-----	1095
Qy	851		GATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAAC	910
Db	1096		-----CCCTGGACACCAAC	1109
Qy	911		TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGG	970
Db	1110		TACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGC	1169
Qy	971		AAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAAGGGCTACCATGCCAATTTCTGCCTG	1030

```

Db      1170 AAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTG 1229
Qy      1031 GGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGGCTCTGTAC 1090
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      1230 GGGCCCTGTCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGGCCCTGTAC 1289
Qy      1091 AACCAGCACAACCCGGGCGCGTCCGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA 1150
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      1290 AACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCA 1349
Qy      1151 CTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1210
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      1350 CTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATC 1409
Qy      1211 GTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCCGGCAG 1270
        ||||| ||||| || ||||||| ||||||| ||||||| || || || || || ||
Db      1410 GTGCGTTCCTGAAAATGCAGCTGAGGCCCCGCCCCGCCCCACCCCGGCAGGCCCG 1469
Qy      1271 GCCCCGCCCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGT 1321
        ||||| ||||||| || || || || || || || || || || || || || || || ||
Db      1470 GCCCCGCCCCACCCACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGT 1529
Qy      1322 GCCCCAAGCCCACTTGGGATCGATTAAA 1349
        ||||||| ||||| || || ||
Db      1530 GCCCCAAGCCCACCTGGGGCCCCATTAA 1557

```

RESULT 15

AAT05876

ID AAT05876 standard; cDNA; 2745 BP.

XX

AC AAT05876;

XX

DT 25-JUN-1996 (first entry)

XX

DE cDNA encoding transforming growth factor-beta 1.

XX

KW macrophage inducible nitric oxide synthase; iNOS; constitutive NOS;

KW interleukin-1-beta; transforming growth factor-beta; TGF-beta; IL1-beta;

KW nitric oxide production; hypotension; inflammation; septic shock;

KW treatment; ds.

XX

OS Mammalian sp.

XX

FH Key Location/Qualifiers

FT CDS 842..2017

FT /\*tag= a

FT /product= transforming growth factor-beta 1

XX

PN WO9526745-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1994; 94WO-US03705.

XX

PR 05-APR-1994; 94WO-US03705.

XX

PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Lee M, Perrella MA;  
 XX  
 DR WPI; 1995-358443/46.  
 DR P-PSDB; AAR83054.  
 XX  
 PT Treatment of hypotension, esp. in septic shock - by administering  
 PT transforming growth factor-beta e.g. to inhibit inducible nitric  
 PT oxide synthase gene transcription  
 XX  
 PS Disclosure; Fig 15; 52pp; English.  
 XX  
 CC The cDNA encodes transforming growth factor-beta 1 (TGF-beta 1) which  
 CC has been found to inhibit inducible nitric oxide synthase (iNOS) gene  
 CC transcription, esp. in interleukin-1-beta (IL1-beta) stimulated rat  
 CC smooth muscle cells, and at a dose which does not inhibit constitutive  
 CC NOS. TGF-beta 1 or 2 (AAR83055) or their active fragments, can be used  
 CC in the treatment of hypotension, such as that associated with severe  
 CC inflammation or septic shock.  
 XX  
 SQ Sequence 2745 BP; 527 A; 938 C; 801 G; 479 T; 0 other;

Query Match 71.2%; Score 969; DB 16; Length 2745;  
 Best Local Similarity 84.7%; Pred. No. 1.3e-186;  
 Matches 1148; Conservative 0; Mismatches 145; Indels 62; Gaps 3;

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    897 TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy     126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy     186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy     246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy     306 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1137 CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy     366 AAATCTATGATAAAATTCAGGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db   1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy     426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT-- 483
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||

```

Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	484	-GAGGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATT 	542
Db	1317	GGAGGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATT	1376
Qy	543	CCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCCGGAGTGGCTGTCTCT 	602
Db	1377	CCTGGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTT	1436
Qy	603	TTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTC 	662
Db	1437	TTGATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTC	1496
Qy	663	GCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAAACACACTCCACGTGGAAATTAACG 	722
Db	1497	GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAAACACACTGCAAGTGGACATCAACG	1556
Qy	723	GGTTCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC 	782
Db	1557	GGTTCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTTC	1616
Qy	783	TGCTCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCC 	842
Db	1617	TGCTTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCC	1676
Qy	843	GAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGG 	902
Db	1677	GA-----GCCCTGG	1685
Qy	903	ATACCAACTACTGCTTCAGCTCCACGGAGAAGAAGTCTGTCGTGCGGCAGCTCTACATTG 	962
Db	1686	ACACCAACTATTGCTTCAGCTCCACGGAGAAGAAGTCTGTCGTGCGGCAGCTGTACATTG	1745
Qy	963	ACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATT 	1022
Db	1746	ACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATT	1805
Qy	1023	TCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGG 	1082
Db	1806	TCTGCCTCGGGCCCTGCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGG	1865
Qy	1083	CTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGC 	1142
Db	1866	CCCTGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGC	1925
Qy	1143	TGGAGCCACTGCCCATCGTGTAACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCA 	1202
Db	1926	TGGAGCCGCTGCCCATCGTGTAACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCA	1985
Qy	1203	ACATGATCGTGCGTTCTTGAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCCA 	1262
Db	1986	ACATGATCGTGCGTCTCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGG	2045
Qy	1263	CCCGGCAGGCCCGGCCCAACCCCGCCCGCCT-----CACCGGGCTGTATTTAAGG 	1314
Db	2046	CAGGCCCGGCCCAACCCCGCCCGCCTGTCCTTGCCCATGGGGCTGTATTTAAGG	2105

```

Qy      1315 ACATCGTGCCCCAAGCCCCTTGGGATCGATTAAA 1349
          ||| ||||| ||||| ||||| ||| | |||
Db      2106 ACACCGTGCCCCAAGCCCACCTGGGGCCCCATTAA 2140

```

Search completed: October 27, 2003, 19:11:34  
Job time : 419.887 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 18:35:27 ; Search time 404.529 Seconds  
(without alignments)  
9022.658 Million cell updates/sec

Title: US-10-017-372E-38  
Perfect score: 1361  
Sequence: 1 tggtagcgagatggcgccctt.....cgattaaagcgccgcgact 1361

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
	1	973	71.5	1821	14	US-10-087-268-1	Sequence 1, Appli
	2	972.6	71.5	2742	14	US-10-037-270-220	Sequence 220, Appl
	3	971.4	71.4	1821	14	US-10-087-268-4	Sequence 4, Appli
	4	969	71.2	2745	11	US-09-948-002-28	Sequence 28, Appl
	5	850.2	62.5	2094	11	US-09-948-002-1	Sequence 1, Appli
	6	830.6	61.0	1585	11	US-09-948-002-27	Sequence 27, Appl
	7	657.2	48.3	1376	10	US-09-756-283A-19	Sequence 19, Appl
	8	594.6	43.7	1352	10	US-09-756-283A-21	Sequence 21, Appl
	9	354.4	26.0	489	11	US-09-911-904-167	Sequence 167, App
	10	308.6	22.7	339	10	US-09-813-271B-1	Sequence 1, Appli
	11	249	18.3	2574	11	US-09-906-158-3	Sequence 3, Appli
	12	249	18.3	2574	13	US-10-028-158-20	Sequence 20, Appl
	13	242.8	17.8	4382	12	US-09-957-458B-9	Sequence 9, Appli
	14	239.6	17.6	2879	11	US-09-906-158-10	Sequence 10, Appl
	15	227.2	16.7	336	10	US-09-813-271B-7	Sequence 7, Appli
	16	209.4	15.4	339	10	US-09-813-271B-5	Sequence 5, Appli
c	17	185.4	13.6	2381	12	US-10-311-455-2384	Sequence 2384, Ap
	18	183	13.4	4267	11	US-09-948-002-47	Sequence 47, Appl
	19	182.4	13.4	336	10	US-09-813-271B-11	Sequence 11, Appl
	20	181.2	13.3	2381	12	US-10-311-455-2383	Sequence 2383, Ap
	21	180.8	13.3	336	10	US-09-813-271B-9	Sequence 9, Appli
	22	158.2	11.6	339	10	US-09-813-271B-3	Sequence 3, Appli
	23	158.2	11.6	2570	12	US-09-960-706-663	Sequence 663, App
	24	158.2	11.6	2912	13	US-10-044-090-323	Sequence 323, App
c	25	141.6	10.4	597	9	US-09-864-761-15319	Sequence 15319, A
c	26	134.8	9.9	154	9	US-09-864-761-31841	Sequence 31841, A
c	27	131.8	9.7	206	10	US-09-833-381-577	Sequence 577, App
c	28	122	9.0	537	9	US-09-864-761-8844	Sequence 8844, Ap
c	29	121	8.9	148	9	US-09-864-761-25510	Sequence 25510, A
	30	117	8.6	181	10	US-09-833-381-1416	Sequence 1416, Ap
	31	104.8	7.7	851	13	US-10-027-632-152938	Sequence 152938,
	32	102.2	7.5	29000	11	US-09-906-158-17	Sequence 17, Appl
c	33	100.4	7.4	224	10	US-09-833-381-73	Sequence 73, Appl
	34	79.4	5.8	450	11	US-09-918-995-6145	Sequence 6145, Ap
c	35	69.8	5.1	431	9	US-09-864-761-18644	Sequence 18644, A
c	36	69.4	5.1	176	10	US-09-833-381-571	Sequence 571, App
	37	69	5.1	584	12	US-10-029-386-9758	Sequence 9758, Ap
	38	68.2	5.0	176	12	US-10-029-386-23458	Sequence 23458, A
c	39	64	4.7	363	10	US-09-833-381-585	Sequence 585, App
	40	62.6	4.6	364	12	US-10-029-386-26476	Sequence 26476, A
	41	62.6	4.6	544	12	US-10-029-386-12776	Sequence 12776, A
	42	62.6	4.6	927	12	US-10-244-718-1	Sequence 1, Appli
	43	57.4	4.2	486	11	US-09-918-995-25641	Sequence 25641, A
	44	55.2	4.1	594	13	US-10-027-632-141376	Sequence 141376,
	45	50.4	3.7	658	9	US-09-765-527-254	Sequence 254, App

#### ALIGNMENTS

US-10-087-268-1

US-10-087-268-1

Matches 1136; Conservative 0; Mismatches 140; Indels 59; Gaps 2;

Db 747 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 806



Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGGCAACC	365
Db	807	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGGCAACG	866
Qy	366	AAATCTATGATAAAATTCAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	867	AAATCTATGACAAGTTCAGGAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	426	AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	545
Db	987	GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	1047	GGCGATACTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTATCTTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTGC	785
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1346	-----GCCCTGGACA	1355
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1356	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1415
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCAAGGGCTACCATGCCAATTTCT	1025
Db	1416	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCT	1475
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1476	GCCTCGGGCCCTGCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1535
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1536	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1595



US-10-037-270-220

Query Match 71.5%; Score 972.6; DB 14; Length 2742;  
Best Local Similarity 84.9%; Pred. No. 2e-254;  
Matches 1149; Conservative 0; Mismatches 144; Indels 60; Gaps 3;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    897 TGGTGCTGACGCCTGGCCGGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy    186 TTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1017 TCGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy    246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 305
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1137 CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy    366 AAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy    426 AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy    486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 1376

Qy    546 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1377 GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG 1436

Qy    606 ATGTCACCGGAGTTGTGCGGCAGTGCGTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC 665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1437 ATGTCACCGGAGTTGTGCGGCAGTGCGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC 1496

Qy    666 TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 725
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1497 TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT 1556

Qy    726 TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGC 785
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1557 TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCTTCCTGC 1616
```

```

Qy      786 TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 845
      |||
Db      1617 TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA- 1675

Qy      846 CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA 905
      |||
Db      1676 -----GCCCTGGACA 1685

Qy      906 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 965
      |||
Db      1686 CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT 1745

Qy      966 TCCGGAAGGACCTGGGCTGGAAGTGGATTATGAAACCAAGGGCTACCATGCCAATTTCT 1025
      |||
Db      1746 TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCT 1805

Qy      1026 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1085
      |||
Db      1806 GCCTCGGGCCCTGCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC 1865

Qy      1086 TGTACAACCAGCACAACCCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1145
      |||
Db      1866 TGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1925

Qy      1146 AGCCACTGCCCATCGTGTA TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205
      |||
Db      1926 AGCCGCTGCCCATCGTGTA TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1985

Qy      1206 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCCGCCCCGCCCCACAGCCCCGCCCCACCC 1265
      |||
Db      1986 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGTCCCCGCCCCGCCCCGCCCCGCCCCGGCAG 2045

Qy      1266 GGCGAGCCCCGGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA 1317
      |||
Db      2046 GCGCGGCCCCACCCCCGCCCCGCCCCGCTGCTTGCCTTGCCCATGGGGGCTGTATTTAAGGACA 2105

Qy      1318 -TCGTGCCCCAAGCCCACCTGGGGATCGATTAAA 1349
      |||
Db      2106 CCGTGCCCCAAGCCCACCTGGGGCCCCATTAA 2138

```

RESULT 3

US-10-087-268-4

; Sequence 4, Application US/10087268

; Publication No. US20030119010A1

; GENERAL INFORMATION:

; APPLICANT: Jonsonn, Julie Ruth

; APPLICANT: Powell, Elizabeth Ellen

; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condition

; FILE REFERENCE: Fibrosis

; CURRENT APPLICATION NUMBER: US/10/087,268

; CURRENT FILING DATE: 2002-03-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

```

; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
US-10-087-268-4

```

```

Query Match          71.4%;  Score 971.4;  DB 14;  Length 1821;
Best Local Similarity 85.0%;  Pred. No. 3.9e-254;
Matches 1135;  Conservative 0;  Mismatches 141;  Indels 59;  Gaps 2;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      507 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 566

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      567 TGGTGCTGACGCCTGGCCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 626

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      627 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 686

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      687 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCCAGGCCGTGCTCG 746

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGAAACCGGAGCCCGAGC 305
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      747 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 806

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      807 CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG 866

Qy      366 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      867 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 926

Qy      426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      927 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 986

Qy      486 GGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      987 GGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT 1046

```

Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACC	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1346	-----GCCCTGGACA	1355
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1356	CCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACT	1415
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCT	1025
Db	1416	TCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCT	1475
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1476	GCCTCGGGCCCTGCCCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC	1535
Qy	1086	TGTACAACCAGCACAAACCGGGCGCGTGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1536	TGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1595
Qy	1146	AGCCACTGCCCATCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1596	AGCCGCTGCCCATCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1655
Qy	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	1656	TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAG	1715
Qy	1266	GGCAGGCCCGGCCCAACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA	1317
Db	1716	GCCCGGCCCAACCCCGCCCGCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACA	1775
Qy	1318	TCGTGCCCCAAGCCC	1332
Db	1776	CCGTGCCCCAAGCC	1790

RESULT 4

US-09-948-002-28

```
; Sequence 28, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 28
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (842)...(2017)
US-09-948-002-28
```

```
Query Match          71.2%;  Score 969;  DB 11;  Length 2745;
Best Local Similarity 84.7%;  Pred. No. 1.9e-253;
Matches 1148;  Conservative 0;  Mismatches 145;  Indels 62;  Gaps 3;
```

```
Qy      6  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      66  TAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      897 TGGTGCTGACGCCTGGCCCGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy     126  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     957  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy     186  TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy     246  CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy     306  CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG 1196

Qy     366  AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db	1197	AAATCTATGACAAGTTC	CAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT--	483	
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316	
Qy	484	-GAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATT	542	
Db	1317	GGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATT	1376	
Qy	543	CCTGGCGCTACCTCAGCAACCGGTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTT	602	
Db	1377	CCTGGCGATACCTCAGCAACCGGTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTT	1436	
Qy	603	TTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTC	662	
Db	1437	TTGATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTC	1496	
Qy	663	GCCTCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACG	722	
Db	1497	GCCTTAGCGCCCACTGCTCTCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACG	1556	
Qy	723	GGTTCAATTCTGGCCGCCGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCC	782	
Db	1557	GGTTCACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCCTTCC	1616	
Qy	783	TGCTCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCC	842	
Db	1617	TGCTTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCC	1676	
Qy	843	GAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGG	902	
Db	1677	GA-----GCCCTGG	1685	
Qy	903	ATACCAACTACTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTG	962	
Db	1686	ACACCAACTATTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTG	1745	
Qy	963	ACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATT	1022	
Db	1746	ACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATT	1805	
Qy	1023	TCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGG	1082	
Db	1806	TCTGCCTCGGGCCCTGCCCTACATTTGAGACCTGGACACGCAGTACAGCAAGGTCCTGG	1865	
Qy	1083	CTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGC	1142	
Db	1866	CCCTGTACAACCAGCATAAACCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGC	1925	
Qy	1143	TGGAGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCA	1202	
Db	1926	TGGAGCCGCTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCA	1985	
Qy	1203	ACATGATCGTGCGTTCCTGCAAGTGACAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCCA	1262	
Db	1986	ACATGATCGTGCGTTCCTGCAAGTGACAGCTGAGGTCCCGCCCCGCCCCGCCCGCCCCGG	2045	



```

Qy      1263 CCCGGCAGGCCCCGGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGG 1314
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2046 CAGGCCCCGGCCCCACCCCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGG 2105

Qy      1315 ACATCGTGCCCCAAGCCCACTTGGGATCGATTAAA 1349
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2106 ACACCGTGCCCCAAGCCCACTTGGGGCCCCATTAA 2140

```

# RESULT 5

US-09-948-002-1

```

; Sequence 1, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 1
; LENGTH: 2094
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (868)...(2040)
US-09-948-002-1

```

```

Query Match          62.5%;  Score 850.2;  DB 11;  Length 2094;
Best Local Similarity 80.9%;  Pred. No. 3.7e-221;
Matches 1040;  Conservative 0;  Mismatches 188;  Indels 57;  Gaps 2;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      863 CCCCCATGCCGCCCTCGGGGCTGCGGCTACTGCCGCTTCTGCTCCCACTCCCGTGGCTTC 922

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      923 TAGTGCTGACGCCCGGGAGGCCAGCCGCGGGACTCTCCACCTGCAAGACCATCGACATGG 982

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      983 AGCTGGTGAACGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAAACTAAGGC 1042

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGGCCCGCTGCCTGAGGCAGTACTGG 245
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1043 TCGCCAGTCCCCCAAGCCAGGGGGAGGTACCGCCCGGGCCCGCTGCCCCAGGCGGTGCTCG 1102

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 305

```

Db	1103	CTTTGTACAAACAGCACCCGCGACCGGGTGGCAGGCGAGAGCGCCGACCCAGAGCCGGAGC	1162
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	1163	CCGAAGCGGACTACTATGCTAAAGAGGTCAACCGCGTGCTAATGGTGGACCGCAACAACG	1222
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1223	CCATCTATGAGAAAACCAAAGACATCTCACACAGTATATATATGTTCTTCAATACGTCAG	1282
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1283	ACATTCCGGGAAGCAGTGCCCGAACCCCATTTGCTGTCCCGTGTCAGAGCTGCGCTTGCA	1342
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	545
Db	1343	GATTAAATCAAGTGTGGAGCAACATGTGGAACTCTACCAGAAATATAGCAACAATTCTT	1402
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	1403	GGCGTTACCTTGGTAACCGGCTGCTGACCCCCACTGATACGCTGAGTGCGTGTCTTTG	1462
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1463	ACGTCACTGGAGTTGTACGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTCGAT	1522
Qy	666	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1523	TCAGCGCTCACTGCTCTTGTGACAGCAAAGATAACAACTCCACGTGGAAATCAACGGGA	1582
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCTGTC	785
Db	1583	TCAGCCCCAAACGTGCGGGCGACCTGGGCACCATCCATGACATGAACCGGCCCTTCTGTC	1642
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1643	TCCTCATGGCCACCCCCCTGGAAAGGGCCAGCACCTGCACAGCTCACGGCACCGGAGA-	1701
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1702	-----GCCCTGGATA	1711
Qy	906	CCAATACTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT	965
Db	1712	CCAATAATTGCTTCAGCTCCACAGAGAAGAACTGCTGTGTGCGGCAGCTGTACATTGACT	1771
Qy	966	TCCGGAAGGACCTGGGCTGGAAAGTGGAATTCATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1772	TTAGGAAGGACCTGGGTGGAAAGTGGAATTCACGAGCCCAAGGGCTACCATGCCAATTTCT	1831
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085
Db	1832	GTCTGGGACCTGCCCCATATTGGAGCCTGGACACACAGTACAGCAAGGTCCTTGCCC	1891
Qy	1086	TGTACAACCAGCACAACCCGGGCGCGTCCGGCGCGCGTGCTGCGTGCCGAGGCGCTGG	1145

Db	1892	TCTACAACCAACACAACCCGGGCGCTTCGGCGTCACCGTGCTGCGTGCCGCAGGCTTTGG	1951
Qy	1146	AGCCACTGCCCATCGTGTA	1205
Db	1952	AGCCACTGCCCATCGTCTACTACGTGGGTCGCAAGCCCAAGGTGGAGCAGTTGTCCAACA	2011
Qy	1206	TGATCGTGCGTTCTCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCC	1265
Db	2012	TGATTGTGCGCTCTCTGCAAGTGCAGCTGAAGCCCCGCCCCG-----CCCGCCCCCTCCC	2065
Qy	1266	GGCAGGCCCCGGCCCCACCCCCGCCC	1290
Db	2066	GGCAGGCCCCGGCCCCGCCCCGCCC	2090

RESULT 6

```

US-09-948-002-27
; Sequence 27, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
;   APPLICANT: Nicholas M. Dean
;   APPLICANT: Susan F. Murray
;   TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
;   TITLE OF INVENTION: FACTOR BETA    EXPRESSION
;   FILE REFERENCE: ISPH-0607
;   CURRENT APPLICATION NUMBER: US/09/948,002
;   CURRENT FILING DATE: 2000-09-05
;   PRIOR APPLICATION NUMBER: 09/661,753
;   PRIOR FILING DATE: 2000-09-14
;   PRIOR APPLICATION NUMBER: 60/154,546
;   PRIOR FILING DATE: 1999-09-17
;   NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 27
;   LENGTH: 1585
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (413)...(1585)
US-09-948-002-27

```

Query Match 61.0%; Score 830.6; DB 11; Length 1585;  
Best Local Similarity 81.3%; Pred. No. 7.4e-216;  
Matches 999; Conservative 0; Mismatches 179; Indels 51; Gaps 1;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	65
Db	408	CCCCATGCCGCCCTCGGGGCTGCGGCTGCTGCCGCTTCTGCTCCCACTCCCGTGGCTTC	467
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	468	TAGTGCTGACGCCCGGAGGCCAGCCGCGGACTCTCCACCTGCAAGACCATCGACATGG	527
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	528	AGCTGGTGAAAACGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAAACTAAGGC	587

Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCCTGAGGCAGTACTGG	245
Db	588	TCGCCAGTCCCCGAGCCAGGGGGAGGTACCGCCGGGCCCGCTGCCCGAGGCGGTGCTCG	647
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	648	CTTTGTACAACAGCACCCGCGACCGGGTGGCAGGCGAGAGCGCTGACCCGGAGCCCGAGC	707
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	708	CCGAGGCGGACTACTACGCCAAAGAAGTCACCCGCGTGCTAATGGTGGACCGCAACAACG	767
Qy	366	AAATCTATGATAAATTCAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	768	CAATCTATGACAAAACCAAAGACATCACACACAGTATATATATGTTCTTCAATACGTCAG	827
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	828	ACATTCCGGGAAGCAGTGCCAGAACCCCCATTGCTGTCCCGTGACAGCTGCGCCTGCAGA	887
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	545
Db	888	GATTCAAGTCAACTGTGGAGCAACACGTAGAACTCTACCAGAAATATAGCAACAATTCCT	947
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	948	GGCGTTACCTTGGTAACCGGCTGCTGACCCCCACTGATACGCCTGAGTGGCTGTCTTTG	1007
Qy	606	ATGTCAACCGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1008	ACGTCACTGGAGTTGTCCGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTTCGCT	1067
Qy	666	TCAGTGCCCACTGTTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1068	TCAGTGCTCACTGCTCTTGTGACAGCAAAGATAATGTACTCCACGTGGAAATCAATGGGA	1127
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1128	TCAGTCCCAAACGTGAGGTGACCTGGGCACCATCCATGACATGAACCGACCTTCTCTGC	1187
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1188	TCCTCATGGCCACCCCCCTGGAAGGGCTCAACACCTGCACAGCTCCAGGCACCGGAGA-	1246
Qy	846	CCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATA	905
Db	1247	-----GCCCTGGATA	1256
Qy	906	CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGTGCGTGCGGCAGCTCTACATTGACT	965
Db	1257	CCAACTACTGCTTCAGCTCCACAGAGAAGAACTGTGCTGTACGGCAGCTGTACATTGACT	1316
Qy	966	TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCT	1025
Db	1317	TTAGGAAGGACCTGGGTTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAATTTCT	1376
Qy	1026	GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC	1085

Db	1377	GTCTGGGGCCCTGCCCTACATTTGGAGCCTGGACACACAGTACAGCAAGGTCCTTGCCC	1436
Qy	1086	TGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1145
Db	1437	TCTACAACCAACACAAACCCGGGTGCTTCCGCATCACCGTGCTGCGTGCCGCAGGCTTTGG	1496
Qy	1146	AGCCACTGCCCATCGTGTA TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1205
Db	1497	AGCCACTGCCCATCGTCTACTACGTGGGTGCGCAAGCCCAAGGTGGAGCAGTTGTCCAACA	1556
Qy	1206	TGATCGTGCGTTCCTGCAAGTGCAGCTGA	1234
Db	1557	TGATCGTGCGCTCCTGCAAGTGCAGCTGA	1585

RESULT 7

US-09-756-283A-19  
; Sequence 19, Application US/09756283A  
; Patent No. US20020151478A1  
; GENERAL INFORMATION:  
; APPLICANT: Chernajovsky, Yuti  
; APPLICANT: Dreja, Hanna Stina  
; APPLICANT: Adams, Gillian  
; TITLE OF INVENTION: Latent Fusion Protein  
; FILE REFERENCE: 0623.1000000  
; CURRENT APPLICATION NUMBER: US/09/756,283A  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 1376  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: LAP-mIFNbeta construct  
; NAME/KEY: CDS  
; LOCATION: (1)..(1368)  
US-09-756-283A-19

Query Match 48.3%; Score 657.2; DB 10; Length 1376;  
Best Local Similarity 88.3%; Pred. No. 1e-168;  
Matches 726; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

Qy	11	ATGGCGCCTTCGGGGCTGCGGCTCTTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	70
Db	1	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	60
Qy	71	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	130
Db	61	CTGACGCCTGGCCCGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	120
Qy	131	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	190
Db	121	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	180
Qy	191	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	250

```

      |||
Db      181 AGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 240
Qy      251 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG 310
      |||
Db      241 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 300
Qy      311 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 370
      |||
Db      301 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 360
Qy      371 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 430
      |||
Db      361 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 420
Qy      431 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT---GAGG 487
      |||
Db      421 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGG 480
Qy      488 CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGG 547
      |||
Db      481 CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG 540
Qy      548 CGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGAT 607
      |||
Db      541 CGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGAT 600
Qy      608 GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTC 667
      |||
Db      601 GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCCTT 660
Qy      668 AGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTC 727
      |||
Db      661 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC 720
Qy      728 AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTTGCTC 787
      |||
Db      721 ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTTGCTT 780
Qy      788 CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGC 829
      |||
Db      781 CTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGC 822

```

RESULT 8

US-09-756-283A-21

; Sequence 21, Application US/09756283A

; Patent No. US20020151478A1

; GENERAL INFORMATION:

; APPLICANT: Chernajovsky, Yuti

; APPLICANT: Dreja, Hanna Stina

; APPLICANT: Adams, Gillian

; TITLE OF INVENTION: Latent Fusion Protein

; FILE REFERENCE: 0623.1000000

; CURRENT APPLICATION NUMBER: US/09/756,283A

; CURRENT FILING DATE: 2001-01-09

; NUMBER OF SEQ ID NOS: 100

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
;   LENGTH: 1352
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: mIFNbeta-LAP construct
;   NAME/KEY: CDS
;   LOCATION: (1)..(1344)
US-09-756-283A-21
```

Query Match 43.7%; Score 594.6; DB 10; Length 1352;  
Best Local Similarity 87.3%; Pred. No. 1.1e-151;  
Matches 664; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

Qy	73	GACGCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGGAGCTGGT	132
Db	582	GGCGGGAGGGGGCTCAGCGGCCGCACTATCCACCTGCAAGACTATCGACATGGAGCTGGT	641
Qy	133	GAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAG	192
Db	642	GAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAG	701
Qy	193	CCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTA	252
Db	702	CCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTA	761
Qy	253	CAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCAGCCAGAGGC	312
Db	762	CAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCAGCCTGAGGC	821
Qy	313	GGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAAGCGGCAACCAAATCTA	372
Db	822	CGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAACCCACAACGAAATCTA	881
Qy	373	TGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCG	432
Db	882	TGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCG	941
Qy	433	GGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT--GAGGCT	489
Db	942	AGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGGCT	1001
Qy	490	CAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCG	549
Db	1002	CAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCG	1061
Qy	550	CTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCCGAGTGCGCTGTCTTTGATGT	609
Db	1062	ATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGTTATCTTTTGATGT	1121
Qy	610	CACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAG	669
Db	1122	CACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGAGGGGAAATTGAGGGCTTTTCGCCTTAG	1181
Qy	670	TGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAA	729

Db 1182 CGCCCACTGCTCCTGTGACAGCAGGGATAACA CACTGCAAGTGGACATCAACGGGTTCAC 1241

Qy 730 TTCTGGCCGCCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCT 789  
| | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1242 TACCGGCCGCCCGAGGTGACCTGGCCACCATT CATGGCATGAACCGGCCCTTCCTGCTTCT 1301

Qy 790 CATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCT 830  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1302 CATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCT 1342

US-09-911-904-167

Query Match 26.0%; Score 354.4; DB 11; Length 489;  
Best Local Similarity 84.1%; Pred. No. 1.8e-86;  
Matches 445; Conservative 0; Mismatches 32; Indels 52; Gaps 2;



[illegible]

## RESULT 10

US-09-813-271B-1

; Sequence 1, Application US/09813271B

GENERAL INFORMATION:

APPLICANT:

(A) Nico Cerletti

TITLE OF INVENTION: New process for the production of  
biologically active protein

```

; NUMBER OF SEQUENCES: 13

```

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. US20020115834A Iartis Patent Department

STREET: 564 Morris Avenue

CITY: Summit

STATE: New Jersey

COUNTRY: USA

ZIP: 07901

COMPUTER READABLE FORM:

```

;
;          MEDIUM TYPE: Floppy disk
;

```

```
COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

```

```

; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/813,271B

FILING DATE: 20-Mar-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP95/02719

FILING DATE: 12-Jul-95

APPLICATION NUMBER: EPO 94810439.3

FILING DATE: 25-Jul-94

ATTORNEY/AGENT INFORMATION:

NAME: Pfeiffer, Hesna J.

REGISTRATION NUMBER: 22640

REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

TELECOMMUNICATION INFORMATION:



; FILE REFERENCE: RTS-0257  
; CURRENT APPLICATION NUMBER: US/09/906,158  
; CURRENT FILING DATE: 2001-07-14  
; NUMBER OF SEQ ID NOS: 168  
; SEQ ID NO 3  
; LENGTH: 2574  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (254)...(1492)  
US-09-906-158-3

Query Match 18.3%; Score 249; DB 11; Length 2574;  
Best Local Similarity 53.5%; Pred. No. 1.2e-57;  
Matches 668; Conservative 0; Mismatches 530; Indels 51; Gaps 5;

```
Qy      36 TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 95
      ||| ||| | | ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      261 TGCACCTGCAAAGGGCTCTGGTGGTCTGGCCCTGCTGAACTTTGCCACGGTCAGCCTCT 320

Qy      96 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 155
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      321 CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 380

Qy     156 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
      || | || ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      381 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440

Qy     216 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
      || ||| ----- ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      441 CCCACG-----TCCCCATACAGGTCTGGCCCTTTACAACAGCACCCGGGAGCTGCTGG 494

Qy     276 CCGGGGAAAGTGTCGAACCGGAGCCCG-----AGCCAGAGGCGGACTACT 320
      || | || || ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      495 AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAACACCGAGTCGGAATACT 554

Qy     321 ACGCCAAGGAGGTCACCCCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 380
      | |||| || || | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      555 ATGCCAAAGAAATCCATAAATTCGACATGATCCAGGGGCTGGCGGAGCACAACGAAGTGG 614

Qy     381 TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 440
      || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      615 CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTTCGCTTCAATGTGTCCTCAGTGGAGA 674

Qy     441 TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAG 500
      | | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      675 AAAATAGAACCAACCTATTCCGAGCAGAATTCCGGGTCTTGCGGTGCCCAACCCAGCT 734

Qy     501 TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG 539
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      735 CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCCAGATCCTTCGGCCAGATGAGCACATTG 794

Qy     540 ATTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGAICTACCGGAGTGGCTGT 599
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      795 CCAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGGCACTGCCGAGTGGCTGT 854
```

Qy 600 CCTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT 659  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 855 CCTTTGATGTCAGTACTGTGCGTGAGTGGCTGTTGAGAAGAGAGTCCAACCTTAGGTC 914

Qy 660 TTCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACA-----CACTCCACGTGG 713  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 915 TAGAAATCAGCATTCACTGTCCATGTCACACCTTTCAGCCCAATGGAGATATCCTGGAAA 974

Qy 714 AAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACC 773  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 975 ACATTACGAGGTGATGGAAATCAAATTCAAAGGCGTGGACAATGAGGATGACCATGGCC 1034

Qy 774 GGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCC 833  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1035 GTGGAGATCTGGGGCGCCT---CAAGAAGCAGAAGGATCACCACAACCCTCATCTAATCC 1091

Qy 834 GGCACCGCCGAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTC 893  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1092 TCATGATGATTCCCCCACACCGGCTCGACAACCCGGGCCAGGGGGGTGAGGAAGAAGC 1151

Qy 894 TGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGC 953  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1152 GGGCTTTGGACACCAATTACTGCTTCCGCAACTTGGAGGAGAAGTCTGTGTGCGCCCCC 1211

Qy 954 TCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACC 1013  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1212 TCTACATTGACTTCCGACAGGATCTGGGCTGGAAGTGGGTCCATGAACCTAAGGGCTACT 1271

Qy 1014 ATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCA 1073  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1272 ATGCCAATTTCTGCTCAGGCCCTTGCCCATACCTCCGAGTGCAGACACAACCCACAGCA 1331

Qy 1074 AGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGC 1133  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1332 CGGTGCTGGGACTGTACAACACTCTGAACCCTGAAGCATCTGCCTCGCCTTGCTGCGTGC 1391

Qy 1134 CGCAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCAAGCCCAAGGTGGAGC 1193  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1392 CCCAGGACCTGGAGCCCCTGACCATCCTGTACTATGTTGGGAGGACCCCCAAAGTGGAGC 1451

Qy 1194 AGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGC 1242  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1452 AGCTCTCCAACATGGTGGTGAAGTCTTGTAATGTAGCTGAGACCCAC 1500

RESULT 12

US-10-028-158-20

; Sequence 20, Application US/10028158

; Publication No. US20020110833A1

; GENERAL INFORMATION:

; APPLICANT: Caniggia, Isabella

; APPLICANT: Post, Martin

; APPLICANT: Lye, Stephen

; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF

; TITLE OF INVENTION: TROPHOBLAST

; FILE REFERENCE: 11757.38USWO

; CURRENT APPLICATION NUMBER: US/10/028,158

```

; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US/09/380,662
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)..(1492)
US-10-028-158-20

```

```

Query Match          18.3%;  Score 249;  DB 13;  Length 2574;
Best Local Similarity 53.5%;  Pred. No. 1.2e-57;
Matches 668;  Conservative 0;  Mismatches 530;  Indels 51;  Gaps 5;

```

```

Qy      36  TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 95
      |||  |||  |  |||  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      261  TGCACCTGCAAAGGGCTCTGGTGGTCTGGCCCTGCTGAACTTTGCCACGGTCAGCCTCT 320

Qy      96  GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 155
      |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      321  CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGGAAGCCA 380

Qy     156  TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
      ||  ||  ||  |||  ||  |||  |||  |||  |||  |||  |||  |||  ||  ||
Db      381  TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440

Qy     216  CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      441  CCCACG-----TCCCCTATCAGGTCTTGCCCTTTACAACAGCACCCGGGAGCTGCTGG 494

Qy     276  CCGGGGAAAGTGTGGAACCGGAGCCCG-----AGCCAGAGGCGGACTACT 320
      |  |  |  |  |  |  |||  |  |  |  |  |  |  |  |  |  |  |  |
Db      495  AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAACACCGAGTCGGAATACT 554

Qy     321  ACGCCAAGGAGGTACCCCGGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 380
      |  ||||  ||  ||  |  |||  |  |  |  |  |  |  |  |  |  |  |
Db      555  ATGCCAAAGAAATCCATAAATTTCGACATGATCCAGGGGCTGGCGGAGCACACGAACCTGG 614

Qy     381  TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 440
      ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      615  CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTTCGCTTCAATGTGTCCTCAGTGAGAGA 674

Qy     441  TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAG 500
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      675  AAAATAGAACCAACCTATTCCGAGCAGAATTCCGGGTCTTGCGGTGCCCAACCCAGCT 734

Qy     501  TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG 539
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      735  CTAAGCGGAATGAGCAGAGGATCGAGCTCTCCAGATCCTTCGGCCAGATGAGCACATTG 794

```









; APPLICANT: Susan M. Freier  
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3  
EXPRESSION  
; FILE REFERENCE: RTS-0257  
; CURRENT APPLICATION NUMBER: US/09/906,158  
; CURRENT FILING DATE: 2001-07-14  
; NUMBER OF SEQ ID NOS: 168  
; SEQ ID NO 10  
; LENGTH: 2879  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (611)...(1843)  
US-09-906-158-10

Query Match 17.6%; Score 239.6; DB 11; Length 2879;  
Best Local Similarity 53.0%; Pred. No. 4.4e-55;  
Matches 658; Conservative 0; Mismatches 539; Indels 45; Gaps 5;

```
QY      36 TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 95
      ||| ||| | | ||| | ||| ||||| | | | |
Db      612 TGCACTTGCAAAGGGCTCTGGTAGTCTGGCCCTGCTGAACTTGGCCACAATCAGCCTCT 671

QY      96 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 155
      ||||| ||| | | | | | | | | | | | | | |
Db      672 CTCTGTCCACTTGCAACAGTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 731

QY     156 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
      || | | ||||| | | ||||| | | ||||| | | | |
Db      732 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCATCGGTGATGA 791

QY     216 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
      | | || | || | | ||||| ||||| ||||| | | |
Db      792 CCCACG-----TCCCTATCAGGTCTGGCACTTTACAACAGCACCCGGGAGTTGCTGG 845

QY     276 CCG-----GGGAAAGTGTCGAACCGGAGCCCGAGCCAGAGGCGGACTACT 320
      | | |||| | | | | | | | | | | | |
Db      846 AAGAGATGCACGGGGAGAGGGAGGAAGGCTGCACTCAGGAGACCTCGGAGTCTGAGTACT 905

QY     321 ACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 380
      | |||| | | | | | | | | | | | | | |
Db      906 ATGCCAAAGAGATCCATAAATTGACATGATCCAGGGACTGGCGGAGCACAATGAACTGG 965

QY     381 TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTCCGGGAAGCGG 440
      | | | | | | | | | | | | | | | | |
Db      966 CCGTCTGCCCCAAAGGAATTACCTCTAAGGTTTTCGTTTCAATGTGTCCTCAGTGGAGA 1025

QY     441 TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCA----- 491
      | | | | | | | | | | | | | | | |
Db     1026 AAAATGGAACCAATCTGTTCCGGGCAGAGTTCGGGTCTTGCGGTGCCCAACCCAGCT 1085

QY     492 -----AGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATG 539
      | | | | | | | | | | | | | | | |
Db     1086 CCAAGCGCACAGAGCAGAGAATTGAGCTCTTCCAGATACTTCGACCGGATGAGCACATAG 1145

QY     540 ATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGAGTGGCTGT 599
```

Db	1146	CCAAGCAGCGCTACATAGGTGGCAAGAATCTGCCCAAGGGGCACCGCTGAATGGCTGT	1205
Qy	600	CCTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT	659
Db	1206	CTTTCGATGTCACTGACACTGTGCGCGAGTGGCTGTTGAGGAGAGAGTCCAACCTGGGTC	1265
Qy	660	TTCGCCTCAGTGCCCACTGTTCCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTA	719
Db	1266	TGGAAATCAGCATCCACTGTCCATGTACACCTTTCAGCCCAATGGAGACATACTGGAAA	1325
Qy	720	ACG-GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCC	778
Db	1326	ATGTTTCATGAGGTGATGGAAATCAAATTCAAAGGAGTGGACAAATGAAGATGACCATGGCC	1385
Qy	779	TTCCTGCTCCTCATGGCCACCCCGTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCAC	838
Db	1386	GTGGAGACCTGGGGCGTCTCAAGAAGCAAAAGGATCACCACAACCCACACCTGATCCTCA	1445
Qy	839	CGCCGAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCT--CTGG	896
Db	1446	TGATGATCCCCCACACCGACTGGACAGCCAGGCCAGGGCAGTCAGAGGAAGAAGAGGG	1505
Qy	897	CCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCT	956
Db	1506	CCCTGGACACCAATTACTGCTTCCGCAACCTGGAGGAGAACTGCTGTGTACGCCCCCTTT	1565
Qy	957	ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATG	1016
Db	1566	ATATTGACTTCCGGCAGGATCTAGGCTGGAAATGGGTCCACGAACCTAAGGGTTACTATG	1625
Qy	1017	CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG	1076
Db	1626	CCAACCTTCTGCTCAGGCCCTTGCCCATACCTCCGCAGCGCAGACACAACCCATAGCACGG	1685
Qy	1077	TCCTGGCTCTGTACAACCAGCACAAACCGGGCGCTCGGCGGCGCGTGCTGCGTGCCGC	1136
Db	1686	TGCTTGGACTATACAACACCTGAACCCAGAGGCGTCTGCCTCGCCATGCTGCGTCCCCC	1745
Qy	1137	AGGCGCTGGAGCCACTGCCCATCGTGTAAGTACGTGGGCGCAAGCCCAAGGTGGAGCAGC	1196
Db	1746	AGGACCTGGAGCCCCTGACCATCTTGTAAGTATGTGGGCAGAAACCCCAAGGTGGAGCAGC	1805
Qy	1197	TGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCC	1238
Db	1806	TGTCCAACATGGTGGTGAAGTCGTGTAAGTGCAGCTGAGGCC	1847

RESULT 15

US-09-813-271B-7

; Sequence 7, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of  
; biologically active protein

```

;      NUMBER OF SEQUENCES: 13
;      CORRESPONDENCE ADDRESS:
;          ADDRESSEE: No. US20020115834Alartis Patent Department
;          STREET: 564 Morris Avenue
;          CITY: Summit
;          STATE: New Jersey
;          COUNTRY: USA
;          ZIP: 07901
;      COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;      CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/813,271B
;          FILING DATE: 20-Mar-2001
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: PCT/EP95/02719
;          FILING DATE: 12-Jul-95
;          APPLICATION NUMBER: EPO 94810439.3
;          FILING DATE: 25-Jul-94
;      ATTORNEY/AGENT INFORMATION:
;          NAME: Pfeiffer, Hesna J. .
;          REGISTRATION NUMBER: 22640
;          REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN
;      TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (908) 522-6940
;          TELEFAX: (908) 522-6955
;      INFORMATION FOR SEQ ID NO: 7:
;          SEQUENCE CHARACTERISTICS:
;              LENGTH: 336 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;          MOLECULE TYPE: other nucleic acid
;          DESCRIPTION: /desc = "recombinant hybrid DNA of
;          IMMEDIATE SOURCE:
;              CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3
;          FEATURE:
;              NAME/KEY: mat_peptide
;              LOCATION: 1..132
;              OTHER INFORMATION: /product= "N-terminal 44 amino
;              acids of human TGF-beta1"
;          FEATURE:
;              NAME/KEY: mat_peptide
;              LOCATION: 133..336
;              OTHER INFORMATION: /product= "C-terminal 68 amino
;              acids of human TGF-beta3"
;          FEATURE:
;              NAME/KEY: CDS
;              LOCATION: 1..336
;              OTHER INFORMATION: /product= "hybrid TGF-beta named
;              TGF-beta1-3"
;          SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-271B-7

```

Query Match

16.7%; Score 227.2; DB 10; Length 336;

Best Local Similarity 79.8%; Pred. No. 6.2e-52;  
Matches 268; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```
Qy      896 GCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTC 955
      |||
Db      1 GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTG 60

Qy      956 TACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCAT 1015
      |||
Db      61 TACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCAT 120

Qy     1016 GCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAG 1075
      |||
Db     121 GCCAATTTCTGCTCAGGCCCTTGCCCATACCTCCGCAGTGCAGACACAACCCACAGCACG 180

Qy     1076 GTCCTGGCTCTGTACAACACAGCACAAACCCGGGCGCGTCCGGCGGCGCGCTGCTGCGTGCCG 1135
      ||
Db     181 GTGCTGGGACTGTACAACACTCTGAACCCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCC 240

Qy     1136 CAGGCGCTGGAGCCACTGCCCATCGTGTAATACTACGTGGGCGCAAGCCCAAGGTGGAGCAG 1195
      |||
Db     241 CAGGACCTGGAGCCCTTGACCATCTGTACTATGTTGGGAGGACCCCCAAAGTGGAGCAG 300

Qy     1196 CTGTCCAACATGATCGTGCGTTTCTGCAAGTGCAGC 1231
      ||
Db     301 CTCTCCAACATGGTGGTGAAGTCTTGTAATGTAGC 336
```

Search completed: October 28, 2003, 09:04:41  
Job time : 411.672 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 10:47:27 ; Search time 3377.39 Seconds  
(without alignments)  
9794.056 Million cell updates/sec

Title: US-10-017-372E-38  
Perfect score: 1361  
Sequence: 1 tggtagcgagatggcgccctt.....cgattaaagcggccgcgact 1361

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

# Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		%						
Result	Query							
No.	Score	Match	Length	DB	ID	Description		
	1	706.8	51.9	1072	12	BM562135	BM562135	AGENCOURT
c	2	701.6	51.6	1201	13	BX355682	BX355682	BX355682
	3	692	50.8	900	13	BX349319	BX349319	BX349319
c	4	690.6	50.7	983	13	BX335351	BX335351	BX335351
c	5	649.2	47.7	1041	13	BX324511	BX324511	BX324511
c	6	610.4	44.8	888	13	BX434425	BX434425	BX434425
	7	610.2	44.8	1043	12	BM555996	BM555996	AGENCOURT
c	8	603.2	44.3	950	13	BX383773	BX383773	BX383773
	9	580.6	42.7	859	9	AL530081	AL530081	AL530081
	10	573.8	42.2	713	10	BE312000	BE312000	601154768
	11	565.8	41.6	717	10	BE260971	BE260971	601153715
	12	565.2	41.5	902	13	BQ675698	BQ675698	AGENCOURT
	13	537.6	39.5	925	12	BI818841	BI818841	603037307
c	14	525	38.6	841	9	AL530080	AL530080	AL530080
c	15	524.6	38.5	956	12	BI084718	BI084718	602869722

16	517	38.0	871	13	BQ952138	BQ952138	AGENCOURT
17	509.8	37.5	1093	12	BQ054305	BQ054305	AGENCOURT
18	506.6	37.2	1013	12	BM469326	BM469326	AGENCOURT
19	504.8	37.1	943	12	BI909079	BI909079	603070060
20	503.4	37.0	773	12	BI195242	BI195242	602944518
21	502	36.9	821	10	BG748049	BG748049	602705544
22	502	36.9	1049	12	BQ053417	BQ053417	AGENCOURT
c 23	496.4	36.5	757	14	CA309731	CA309731	UI-H-FT1-
c 24	489.2	35.9	1047	13	BX325288	BX325288	BX325288
25	476	35.0	675	10	BE261764	BE261764	601147839
c 26	469.4	34.5	727	9	AI131171	AI131171	qc15c11.x
c 27	464.2	34.1	785	9	AI148173	AI148173	qb56d01.x
28	464.2	34.1	960	13	BU957001	BU957001	AGENCOURT
29	463	34.0	637	12	BG938856	BG938856	cn28g07.y
30	460.4	33.8	565	14	CD287835	CD287835	l_M15.abd
31	459	33.7	748	10	BF682938	BF682938	602117568
32	458.6	33.7	814	12	BI686959	BI686959	603313837
c 33	453.8	33.3	722	13	BU632686	BU632686	UI-H-FE1-
34	452.2	33.2	548	12	BM068788	BM068788	ie85a02.y
c 35	452	33.2	928	13	BX412434	BX412434	BX412434
c 36	451.8	33.2	695	14	CA425775	CA425775	UI-H-FE1-
37	450	33.1	982	13	BQ672677	BQ672677	AGENCOURT
c 38	449.8	33.0	697	13	BU633535	BU633535	UI-H-FL1-
39	449.2	33.0	659	10	BG500488	BG500488	602544826
c 40	449.2	33.0	700	13	BU633922	BU633922	UI-H-FL1-
c 41	448.6	33.0	690	13	BU625377	BU625377	UI-H-FG1-
c 42	448.6	33.0	697	14	CA426391	CA426391	UI-H-FE1-
c 43	447.8	32.9	778	9	AI743724	AI743724	wg53b06.x
44	447.6	32.9	981	12	BM802192	BM802192	AGENCOURT
c 45	446.8	32.8	1015	10	BG683840	BG683840	602651694

# ALIGNMENTS

## RESULT 1

BM562135

LOCUS BM562135 1072 bp mRNA linear EST 20-FEB-2002

DEFINITION AGENCOURT\_6562032 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5745463  
5', mRNA sequence.

ACCESSION BM562135

VERSION BM562135.1 GI:18807966

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1072)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation



```

Db      380 GCTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTG 439
Qy      547 GCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGA 606
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      440 GCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGA 499
Qy      607 TGTCAACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCT 666
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      500 TGTCAACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCT 559
Qy      667 CAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTT 726
      || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      560 TAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTT 619
Qy      727 CAATTCTGGCCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTCT 786
      || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      620 CACTACCGGCCCGGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGTCT 679
Qy      787 CCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGC 846
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      680 TCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGC 739
Qy      847 CCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATAC 906
      ||||| |||
Db      740 -----CCCTGGACAC 749
Qy      907 CAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTT 966
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      750 CAACTATTGCTTCAGCTCCCGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTT 809
Qy      967 CCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTG 1026
      ||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      810 CCGCAGGGACCTCGGCTGGAAGTGGATCCACGAGCCCCAAGGCTACCATGCCCACTTCTG 869
Qy      1027 CCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCT 1086
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      870 CCTCGGGCCCTGCCCTACATTTGGAGCCTGGACACGCGTACCAGCAAGTCCCTGGCCCT 929
Qy      1087 GTACAACCAGCACAAACCGGGCGCGTCCGGCGGCGCCGTGCTGC--GTGCCGAGGCGCTG 1144
      ||||| || || ||||| || ||||| ||||| ||||| |||||
Db      930 GTACACCCGCCATAAACCGGGCGCCCTCGGCGGCGCGGGTTTGCTGGGCCCCAGGCCCTG 989
Qy      1145 GAGCCACTGCCCAT--CGTGTAACGTGGGCGCG--AAGCCCAAGGTGGAGCAGCTGTC 1200
      || || ||||| || ||||| ||||| ||||| ||||| |||||
Db      990 GACCCCTGCCCCTTCGGGGTACTACCTGGGGCGCCAAGCCCAAGTGGGAACACCTGTC 1049
Qy      1201 CAACAT 1206
      | ||||
Db      1050 CCACAT 1055

```

# RESULT 2

BX355682/c

LOCUS BX355682 1201 bp mRNA linear EST 05-MAY-2003  
 DEFINITION BX355682 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0DI002YJ05 3-PRIME, mRNA sequence.  
 ACCESSION BX355682



VERSION BX355682.1 GI:30371987  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9160.r For  
 more information about this cluster, see  
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI002CE03NP1&cluster=9160.r)  
[cgi-bin/cluster.cgi?seq=CS0DI002CE03NP1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI002CE03NP1&cluster=9160.r). Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DI002CE03NP1.  
 FEATURES Location/Qualifiers  
 source 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DI002YJ05"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 198 a 326 c 386 g 253 t 38 others  
 ORIGIN

Query Match 51.6%; Score 701.6; DB 13; Length 1201;  
 Best Local Similarity 82.1%; Pred. No. 5.1e-143;  
 Matches 904; Conservative 8; Mismatches 124; Indels 65; Gaps 7;

```

Qy      260 ACCCGCGACCGGGTAGCCGGGGAAAGTGTCTGAACCGGAGCCCGAGCCAGAG-GCGGACTA 318
      | ||||| ||||| ||||| | :| ||||| ||||| ||| ||: |||
Db     1046 AACCGCGACCGGGTGGCCGGG--AGAKKAGAAACCGGAGCCCGAGCCTGAGCCCGRACTA 989

Qy      319 CTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAA 378
      ||||| ||||| : ||||| ||||| ||| ||||| |||
Db     988 CTACGCCAAGGAGGTMAACCCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATGACAA 929

Qy      379 ATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGC 438
      ||||| | ||||| ||||| ||||| ||||| ||||| |||||
Db     928 GTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGC 869

Qy      439 GGTGCCGGAACCTGTATTGCTCT-CTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAA 497
      ||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db     868 GGTACCTGAACCCGTGTTGCTCTCCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAA 809
  
```

Qy 498 AAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCA 557  
 |||  
 Db 808 AAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCA 749

Qy 558 GCAACCGGCTGCTGGCCCCAGTGA CTACCCGAGTGGCTGTCTTTGATGTCACCGGAG 617  
 |||  
 Db 748 GCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTCACCGGAG 689

Qy 618 TTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACT 677  
 |||  
 Db 688 TTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGA-GGCTTTCGCCTTAGCGCCCACT 630

Qy 678 GTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCC 737  
 |  
 Db 629 GCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGA CATCAACGGGTTCACTACCGGCC 570

Qy 738 GCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTGCTCCTCATGGCCA 797  
 |||  
 Db 569 GCCGAGGTGACCTGGCCACCATT CATGGCATGAACCGGCCTTCTGCTTCTCATGGCCA 510

Qy 798 CCCCCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCA 857  
 |||  
 Db 509 CCCCCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC---- 454

Qy 858 ACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGCT 917  
 |||  
 Db 453 -----ACCAACTATTGCT 441

Qy 918 TCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 977  
 |||  
 Db 440 TCAGCTCCACGAGAAGNACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC 381

Qy 978 TGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1037  
 |  
 Db 380 TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCA ACTTCTGCCTCGGGCCCT 321

Qy 1038 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1097  
 |  
 Db 320 GCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 261

Qy 1098 ACAACCCGGGCGCGTCCGGCGGCCGTGCTGCGTGCCG CAGGCGCTGGAGCCACTGCCCA 1157  
 |||  
 Db 260 ATAACCCGGGCGCCTCGGCGGCCGTGCTGCGTGCCG CAGGCGCTGGAGCCGCTGCCCA 201

Qy 1158 TCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1217  
 |||  
 Db 200 TCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 141

Qy 1218 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCCCACCCGGCAGGCCCCGGC 1277  
 |||  
 Db 140 CCTGCAAGTGCAGCTGAGGTCCCCGCCCCGCCCGCCCCGCCCCRGAGGCCCCGGCCCCAC 81

Qy 1278 CCCACCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG 1329  
 |||  
 Db 80 CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAA-GACACCCGTCCCCAAG 22

Qy 1330 CCCACTTGGGATCGATTAAAG 1350

Db 21 ||||| ::||:||||  
CCCACNNAMNNMMDATHAAAS 1

### RESULT 3

BX349319

LOCUS	BX349319	900 bp	mRNA	linear	EST 05-MAY-2003
-------	----------	--------	------	--------	-----------------

DEFINITION BX349319 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED

Homo sapiens cDNA clone CS0DL010YL07 5-PRIME, mRNA sequence.

ACCESSION BX349319

VERSION BX349319.1 GI:30379410

KEYWORDS EST.

SOURCE	Homo sapiens (human)
--------	----------------------

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 900)

AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT                      Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

```
cgi-bin/cluster.cgi?seq=CS0BAG059ZD04 CS05596 1&cluster=9160.r.
```

Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAG059ZD04 CS05596 1.

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .900
```

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="CS0DL010YL07"
```

```
/cell type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
```

```

/cell line="RAMOS CELL LINE"

```

```
/clone lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
```

25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

BASE COUNT      179 a      292 c      270 g      155 t      4 others

Query Match 50.8%; Score 692; DB 13; Length 900;

Best Local Similarity 88.1%; Pred. No. 5.8e-141;

Matches 752; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65

[illegible]

Db 30 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 89

Qy	66	TAGTGCTGACGCCTGGCCCGGCCGCGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	90	TGGTGTCTGACGCCTGGCCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG	149
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	150	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	209
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	210	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	269
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGAGCCCGAGC	305
Db	270	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGAGCCCGAGC	329
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	330	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG	389
Qy	366	AAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	425
Db	390	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	449
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	450	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	509
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	510	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	569
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCTTTTG	605
Db	570	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTATCTTTTG	629
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGCGTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	630	ATGTCACCGGAGTTGTGCGGCAGTGCTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	689
Qy	666	TCAGTGCCC ACTGTT CCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	690	TTAGCGCCC ACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	749
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	750	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATT CATGGCATGAACCGGCCCTTCTCTGC	809
Qy	786	TCCTCATGGCCACCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	810	TTCTCATGGCCACCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGNCGAG	869
Qy	846	CCCTGGATACCAAC	859
Db	870	CCCTGGNACACCAC	883



Qy	458	CTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAG	517
Db	803	CTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAG	744
Qy	518	CTATACCAGAAATACAGCAATGATTCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC	577
Db	743	CTGTACCAGAAATACAGCAACAATYCTTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	684
Qy	578	AGTGACTCACCGGAGTGGCTGTCTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACC	637
Db	683	AGCGACTCGCCAGAGTGGTTATCTTTTGTATGTACCGGAGTTGTGCGGCAGTGGTTGAGC	624
Qy	638	CGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGAT	697
Db	623	CGTGGAGGGGAAATTGA-GGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	565
Qy	698	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	757
Db	564	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	505
Qy	758	ATTCACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAG	817
Db	504	ATTCATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAG	445
Qy	818	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACAGCTACCCATACGACGTG	877
Db	444	CATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC-----	409
Qy	878	CCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAAC	937
Db	408	-----ACCAACTATTGCTTCAGCTCCACGGAGAAGAAC	376
Qy	938	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTAT	997
Db	375	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	316
Qy	998	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1057
Db	315	GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTG	256
Qy	1058	GACACTCAGTACAGCAAGGTCTTGGCTCTGTACAACCAGCACAAACCGGGCGCGTGGGCG	1117
Db	255	GACACGCAGTAMMMCAAGGTCTTGGCCMTGTACAACCAGCATAACCC-GGCGCCTCGGCG	197
Qy	1118	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1177
Db	196	GCGCCGTNGT-CGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCTGC	138
Qy	1178	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGAGGC	1237
Db	137	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTTGCAAGTGCAGCTGAGGT	78
Qy	1238	CCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGCCCAACCCCGCCCCGCCCTCAC	1297
Db	77	CCCGCCCCGCC-----CCGCCCCGCCCGGCAGGCCCGGCCCAACCCCGCCCCGCCCGCG	23

Qy 1298 CGGGGCTG 1305  
| | |  
Db 22 CTGCCTTG 15

## RESULT 5

BX324511/c

LOCUS BX324511 1041 bp mRNA linear EST 02-MAY-2003

DEFINITION BX324511 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DC024YD20 3-PRIME, mRNA sequence.

ACCESSION BX324511

VERSION BX324511.1 GI:30332381

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1041)

AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL      Unpublished

COMMENT            Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

cgi-bin/cluster.cgi?seq=CS0AC024DB10NP2&cluster=9160.r. Contact :

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AC024DB10NP2.

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .1041
```

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="CS0DC024YD20"
```

```
/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"
```

```
/clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
```

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT      176 a      283 c      346 g      219 t      17 others

## ORIGIN

Query Match 47.7%; Score 649.2; DB 13; Length 1041:

Best Local Similarity 80.6%; Pred. No. 1.4e-131;

Matches 870; Conservative 10; Mismatches 133; Indels 67; Gaps 8;

Qy 279 GGGAAAGTGTCGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCC 338

: || | : | ||| ||| ||| ||| ||| ||| ||| : | : | : || | | : ||| ||| |||

Db 1022 SGGGAGAKGSAGAACCGGAGCCCGACCCCTRASCCT--ACTAYACCCCAAGRAGTCACCC 966

Qy 339 GCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCACA 398

Db	965	GCCTGYTAATGKT-GAAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACACACA	907
Qy	399	GCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGC	458
Db	906	GCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGTTGC	847
Qy	459	TCTCTC-GGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG	517
Db	846	TCTCCCSGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG	787
Qy	518	CTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC	577
Db	786	CTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	727
Qy	578	AGTGACTCACCGGAGTGGCTGTCTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACC	637
Db	726	AGCGACTCGCCAGAGTGGTTATCTTTTGATGTACCGGAGTTGTGCGGCAGTGGTTGAGC	667
Qy	638	CGCAGAGAGGCTATAGAGGGTTTTGCGCTCAGTGCCCACTGTTCTGTGACAGCAAAGAT	697
Db	666	CGTGGAGGGGAAATTGA-GGCTTTCGCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	608
Qy	698	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	757
Db	607	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	548
Qy	758	ATTCACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCAG	817
Db	547	ATTCATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCGCTGGAGAGGGCCAG	488
Qy	818	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACAGCTACCCATACGACGTG	877
Db	487	CATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC-----	452
Qy	878	CCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAAC	937
Db	451	-----ACCAACTATTGCTTCAGCTCCACGGAGAAGAAC	419
Qy	938	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCT	997
Db	418	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	359
Qy	998	GAACCCAAGGGCTACCATGCCAATTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1057
Db	358	GAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTG	299
Qy	1058	GACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCCGGCG	1117
Db	298	GACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC-GGCGCCTCGGCG	240
Qy	1118	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1177
Db	239	GCGCCGTG-TGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	181
Qy	1178	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGC	1237



```

Db      180 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGAGGT 121
Qy      1238 CCCGCCCCGCCCACAGCCCCGCCCCACCCGGCAGGCCCGCCCCACCCCGCCCGCCT--- 1294
          ||||| ||||| || || || || || || || || || || || || || || || || ||
Db      120 CCCGCCCCGCCCCGCCCCGCCCCGGCAGGCCCGCCCCACCCCGCCCCGCCCCGCTGCC 61
Qy      1295 -----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTGGGATCGATTAAA 1349
          || ||||| ||||| ||||| || || || || || || || || || || || || || || ||
Db      60 TTGCCCATGGGGGCTGTATTTAAGGACACCCGTCCCCAAGCCCACTGGSCACCCAYTAA 1

```

# RESULT 6

BX434425/c

LOCUS BX434425 888 bp mRNA linear EST 15-MAY-2003

DEFINITION BX434425 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YE16  
3-PRIME, mRNA sequence.

ACCESSION BX434425

VERSION BX434425.1 GI:30779291

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 888)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For  
more information about this cluster, see

<http://www.genoscope.cns.fr/>

[cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r). Contact :

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Paraday Avenue Genoscope sequence ID : CS0BAK028AB08NM1.

## FEATURES

source

Location/Qualifiers

1..888

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DE014YE16"

/tissue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT\_6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

BASE COUNT 156 a 253 c 302 g 176 t 1 others

ORIGIN

Query Match 44.8%; Score 610.4; DB 13; Length 888;

Best Local Similarity 82.6%; Pred. No. 4e-123;

Matches 772; Conservative 0; Mismatches 102; Indels 61; Gaps 4;

```
Qy      428 CTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGG 487
      ||||| ||||| || ||||| | ||| | ||||| ||||| ||||| ||
Db      888 CTCCAAAAAGCGGTACCTGAACCTTGTTGTCTNCCCGGCAGAGCTGCGTCTGCTGA-G 830

Qy      488 CTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTTG 547
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      829 CTCAAGTTTAAAAATGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTTG 770

Qy      548 CGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGA CTACCGGAGTGGCTGTCCTTTGAT 607
      || ||||| ||||| || ||||| || ||||| || ||||| ||
Db      769 CGATACCTCAGCAACCGGCTGCTGGCACACAGCGACTCGCCAGAGTGGTTATCTTTTGAT 710

Qy      608 GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTC 667
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      709 GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT 650

Qy      668 AGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTC 727
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      649 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC 590

Qy      728 AATTCTGGCCGCCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGCTC 787
      || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      589 ACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCCTTCCTGCTT 530

Qy      788 CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 847
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      529 CTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA--- 473

Qy      848 CTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACC 907
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      472 -----GCCCTGGACACC 461

Qy      908 AACTACTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTC 967
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      460 AACTATTGCTTCAGCTCCACGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTC 401

Qy      968 CGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGC 1027
      || ||||| ||||| ||||| ||||| || ||||| ||||| ||||| ||
Db      400 CGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGC 341

Qy      1028 CTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTG 1087
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      340 CTCGGGCCCTGCCCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTG 281

Qy      1088 TACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAG 1147
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      280 TACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAG 221

Qy      1148 CCACTGCCCATCGTGTA CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATG 1207
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      220 CCGTGCCCATCGTGTA CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATG 161

Qy      1208 ATCGTGCGTTCTCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCAAGCCCCGCCCCGCGG 1267
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      160 ATCGTGCGTCTCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCCGCCCGCCCCGCGG 101
```

```

Qy      1268 CAGGCCCCGCCCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA-T 1318
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      100 CCGGCCCCACCCCCGCCCCGCCCCGCTGCCTTGCCCATGGGGCTGTATTTAAGGACACC 41

Qy      1319 CGTGCCCCAAGCCCACCTTGGGATCGATTAAAGCGG 1353
          | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      40 CGTGCCCCAAGCCCACCTGGGCCCCATTAAAGAGG 6

```

## RESULT 7

BM555996

LOCUS	BM555996	1043 bp	mRNA	linear	EST 20-FEB-2002
-------	----------	---------	------	--------	-----------------

DEFINITION AGENCOURT\_6544437 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5550039  
5', mRNA sequence.

ACCESSION BM555996

VERSION BM555996.1 GI:18796907

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1043)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT      Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LAM12261 row: i column: 16

High quality sequence stop: 702.

## FEATURES

Location/Qualifiers

source

1. .1043

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="IMAGE:5550039"
```

```
/tissue type="duodenal adenocarcinoma, cell line"
```

```
/lab host="DH10B (phage-resistant)"
```

```
/clone lib="NIH MGC 88"
```

```
/note="Organ: small intestine; Vector: pCMV-SPORT6;
```

Site 1: NotI; Site 2: SalI; Cloned unidirectionally;

oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life

Technologies. Note: this is a NIH MGC Library."

BASE COUNT	202 a	329 c	345 g	154 t	13 others
------------	-------	-------	-------	-------	-----------

## ORIGIN

Query Match 44.8%; Score 610.2; DB 12; Length 1043;

Best Local Similarity 81.1%; Pred. No. 4.7e-123;

Matches 736; Conservative 0; Mismatches 120; Indels 51; Gaps 1;

Qy	396	ACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTAT	455
Db	1	ACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGT	60
Qy	456	TGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG	515
Db	61	TGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG	120
Qy	516	AGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCC	575
Db	121	AGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCAC	180
Qy	576	CCAGTGACTCACCGGAGTGGCTGTCTTTGATGTCAACCGGAGTTGTGCGGCAGTGGCTGA	635
Db	181	CCAGCGACTCGCCAGAGTGGTTATCTTTTGTGATGTCAACCGGAGTTGTGCGGCAGTGGTTGA	240
Qy	636	CCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAG	695
Db	241	GCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG	300
Qy	696	ATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCA	755
Db	301	ATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCA	360
Qy	756	CCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCC	815
Db	361	CCATTGATGGCATGAACCGGCCCTTCCTGCTTCTCATGGCCACCCGCTGGAGAGGGCCC	420
Qy	816	AGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACAGCTACCCATACGACG	875
Db	421	AGCATCTGCAAAGCTCCCGGCACCGCCGA-----	449
Qy	876	TGCCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGA	935
Db	450	-----GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGA	489
Qy	936	ACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTC	995
Db	490	ACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCC	549
Qy	996	ATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCC	1055
Db	550	ACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCC	609
Qy	1056	TAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGG	1115
Db	610	TGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGG	669
Qy	1116	CGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCC	1175
Db	670	CGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCC	729
Qy	1176	GCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTGCAAGTGCAGCTGAG	1235
Db	730	GCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAGTGACCTTGAA	789

Qy 1236 GCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGCCCAACCCCGCCCGCCTC 1295  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 790 GGGTCCGGCCCCCACNCGGNCNNNNNGGNNNNGGGGGNNCCCCGGCGCCGGGGCCCCCGC 849

Qy 1296 ACCGGGG 1302  
 | | |  
 Db 850 GGCCCCG 856

RESULT 8

BX383773/c

LOCUS	BX383773	950 bp	mRNA	linear	EST 08-MAY-2003
-------	----------	--------	------	--------	-----------------

DEFINITION BX383773 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens cDNA clone CS0DK001YA15 3-PRIME, mRNA sequence.

ACCESSION BX383773

VERSION BX383773.1 GI:30457168

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 950)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL      Unpublished

COMMENT            Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

cgi-bin/cluster.cgi?seq=CS0DK001AA08NP1&cluster=9160.r. Contact :

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DK001AA08NP1.

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .950
```

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="CS0DK001YA15"
```

```
/cell type="HELA CELLS COT 25-NORMALIZED"
```

```
/cell_line="HELA"
```

```
/clone lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
```

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT	171 a	260 c	312 g	195 t	12 others
------------	-------	-------	-------	-------	-----------

## ORIGIN

Query Match 44.3%; Score 603.2; DB 13; Length 950;

Best Local Similarity 79.4%; Pred. No. 1.5e-121;

Matches 791; Conservative 8; Mismatches 134; Indels 63; Gaps 5;

Qy	353	GAAAGCGGCAACCAAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTG	412
Db	943	GRAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGMACACACAGCATATATATGTTC	884
Qy	413	TTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAG	472
Db	883	TTCAAMACATCAGAGCTCCGAGAAGCGGTACCTGVACCCGTGTTGCTCTCCCGGGCAGAG	824
Qy	473	CTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATAC	532
Db	823	CTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATAC	764
Qy	533	AGCAATGATTCTGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAG	592
Db	763	AGCAACAATAMCTGGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAG	704
Qy	593	TGGCTGTCCTTTTGATGTACCCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATA	652
Db	703	TGGTTATCTTTTGATGTACCCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATT	644
Qy	653	GAGGGTTTTTCGCCTCAGTGCCCACTGTTCTGTGTACAGCAAAGATAACACACTCCACGTG	712
Db	643	GA-GGCTTTTCGCCTTAGCGCCCACTGCTCCTGTTACAGCAGGGATTACACACTGCAAGTG	585
Qy	713	GAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAAC	772
Db	584	GACATCAACGGGTTTACTACCGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAAC	525
Qy	773	CGGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCC	832
Db	524	CGGCCTTTCCTGCTTCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCC	465
Qy	833	CGGCACCGCCGAGCCCTGGATACCAACAGCTACCCATACGACGTGCCAGACTACGCATCT	892
Db	464	CGGCACCGCCGA-----	453
Qy	893	CTGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAG	952
Db	452	---GCMCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAG	396
Qy	953	CTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTAC	1012
Db	395	CTGTACATTKACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTAC	336
Qy	1013	CATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGC	1072
Db	335	CATGCCGACTTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTGGACACGCAGTACAGC	276
Qy	1073	AAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTG	1132
Db	275	AAGGTCCTGGCGCCTGTACAACCAGCATAACCGGGCGCCTCGGCGGCGCCGTG-TGCGTG	217
Qy	1133	CCGCAGGCGCTGGAGCCACTGCCCATCGTGTAATACTGCGGGCCGCAAGCCCAAGGTGGAG	1192
Db	216	CCGCAGGCGCTGGAGCCGCTGCCCATCGTGTAATACTGCGGGCKKCAAGCCCAAGGTGGAG	157
Qy	1193	CAGCTGTCCAACATGATCGTGCCTTCTGCAAGTGAGCTGAGGCCCCGCCCCGCCCA	1252

Db 156 CAGCTGTCCNACATGATCGTGCGCTCCTGCAAGTGCAGCTGAGATCCCGCCCCGCCCGC 97  
 Qy 1253 GCCCCGCCCAACCGGCAGGCCCGGCCCAACCCCGCCCGC-----CTCACCGGGGC 1303  
 Db 96 CCCGCCCGGAAGGCCGCGGCCCAACCCGCCCGCCCCSCTGCCTTGCCCATGGGGGC 37  
 Qy 1304 TGTATTTAAGGA-CATCGTGCCCCAAGCCCACTTGG 1338  
 Db 36 TGTATTTAAGNACCCCCCTGGCCCAAGCCCACTTGG 1

## RESULT 9

AL530081

LOCUS	AL530081	859 bp	mRNA	linear	EST 23-MAY-2003
-------	----------	--------	------	--------	-----------------

DEFINITION AL530081 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens cDNA clone CS0DD009YM06 5-PRIME, mRNA sequence.

ACCESSION AL530081

VERSION AL530081.2 GI:31067916

KEYWORDS EST.

SOURCE	Homo sapiens (human)
--------	----------------------

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 859)

AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT      On Feb 13, 2001 this sequence version replaced gi:12793574.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

cgi-bin/cluster.cgi?seq=CS0DD009BG03QP1&cluster=9160.r. Contact :

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD009BG03QP1.

FEATURES	Location/Qualifiers
----------	---------------------

source

1.859

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="CS0DD009YM06"
```

```
/tissue type="NEUROBLASTOMA COT 50-NORMALIZED"
```

```
/clone lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
```

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT	179 a	272 c	257 g	148 t	3 others
------------	-------	-------	-------	-------	----------

ORIGIN

Query Match

42.7%; Score 580.6; DB 9; Length 859;

Best Local Similarity 87.3%; Pred. No. 1.3e-116;  
Matches 645; Conservative 2; Mismatches 91; Indels 1; Gaps 1;

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     121 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 180

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     181 TGGTGCTGACGCCTGGCCGGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGG 240

Qy     126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     241 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 300

Qy     186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     301 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCAGGCGCTGCTCG 360

Qy     246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     361 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 420

Qy     306 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCCGCTGCTAATGGTGGAAGCGGCAACC 365
      |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     421 CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCTGCTAATGGTGGAACCCACAACG 480

Qy     366 AAATCTATGATAAAATCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     481 AAATCTATGACAAGTTCAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 540

Qy     426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     541 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 600

Qy     486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTTCCT 545
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     601 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTTCCT 660

Qy     546 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTG-GCTGTCCTTT 604
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     661 GCGGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTGTTATCTTTT 720

Qy     605 GATGTCACCGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGC 664
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     721 GATGTCACCGAGTTGTGCGGCAGTKGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGC 780

Qy     665 CTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGG 724
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     781 CTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGRACATCAACGGG 840

Qy     725 TTCAATTCTGGCCCGCGG 743
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     841 TTCCTACCGGCCCGCGAG 859

```



BE312000  
 LOCUS BE312000 713 bp mRNA linear EST 26-OCT-2000  
 DEFINITION 601154768F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3510592 5',  
 mRNA sequence.  
 ACCESSION BE312000  
 VERSION BE312000.1 GI:9130128  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 713)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LLCM195 row: h column: 17  
 High quality sequence start: 2  
 High quality sequence stop: 713.  
 FEATURES  
 source Location/Qualifiers  
 1..713  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3510592"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_19"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 157 a 226 c 211 g 119 t  
 ORIGIN

Query Match 42.2%; Score 573.8; DB 10; Length 713;  
 Best Local Similarity 87.8%; Pred. No. 3.7e-115;  
 Matches 626; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 135 AGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCC 194  
 |||  
 Db 1 AGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCC 60  
 Qy 195 CCCCAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTTTACA 254  
 |||  
 Db 61 CCCCAGCCAGGGGGAGGTGCCGCCCCGGCCCCGCTGCCCGAGGCCGTGCTCGCCCTGTACA 120

Qy	255	ACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAGGCGG	314
Db	121	ACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGAACCAGGAGCCCGAGCCTGAGGCCG	180
Qy	315	ACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATG	374
Db	181	ACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATG	240
Qy	375	ATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGG	434
Db	241	ACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAG	300
Qy	435	AAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGT	494
Db	301	AAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGT	360
Qy	495	TAAAAGTGGAGCAGCACGTCGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGCTACC	554
Db	361	TAAAAGTGGAGCAGCACGTCGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGATACC	420
Qy	555	TCAGCAACCGGCTGCTGGCCCCCAGTGACTACCCGAGTGGCTGTCTTTGATGTCAACG	614
Db	421	TCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTCAACG	480
Qy	615	GAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCC	674
Db	481	GAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCC	540
Qy	675	ACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAATTAACGGGTTCAATTCTG	734
Db	541	ACTGCTCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCG	600
Qy	735	GCCGCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGCTCCTCATGG	794
Db	601	GCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCCTTCCTGCTTCTCATGG	660
Qy	795	CCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC	847
Db	661	CCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCC	713

BE260971

JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLCM193 row: m column: 04  
High quality sequence stop: 713.

```

Db      300 AGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTT 359
Qy      496 AAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGCTACCT 555
      |||
Db      360 AAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGGATACCT 419
Qy      556 CAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGTCACCGG 615
      |||
Db      420 CAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTGATGTCACCGG 479
Qy      616 AGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCA 675
      |||
Db      480 AGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCCTTAGCGCCCA 539
Qy      676 CTGTTTCTGTGACAGCAAAGATAACACACTCCCGTGGAATTAACGGGGTTCAATTCTGG 735
      |||
Db      540 CTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGGTTCACTACCGG 599
Qy      736 CCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGC 795
      |||
Db      600 CCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTCATGGC 659
Qy      796 CACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGA 852
      |||
Db      660 CACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCCCTGGGA 716

```

# RESULT 12

BQ675698

LOCUS BQ675698 902 bp mRNA linear EST 15-JUL-2002

DEFINITION AGENCOURT\_8036532 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6211917  
5', mRNA sequence.

ACCESSION BQ675698

VERSION BQ675698.1 GI:21786532

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 902)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2375 row: c column: 22

High quality sequence stop: 599.

FEATURES Location/Qualifiers

source 1..902

/organism="Homo sapiens"

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6211917"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_102"  
/note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT        194 a        292 c        255 g        161 t  
ORIGIN

Query Match                    41.5%;    Score 565.2;    DB 13;    Length 902;  
Best Local Similarity        87.0%;    Pred. No. 3.1e-113;  
Matches 621;    Conservative        0;    Mismatches    93;    Indels        0;    Gaps        0;

```
Qy      156  TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1   TCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCCCCGAGCCAGGGGGAGGTGC 60

Qy      216  CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       61  CGCCCGGCCCGCTGCCCAGGCGCTGCTCGCCCTGTACAACAGCACCCGCGACCGGGTGG 120

Qy      276  CCGGGGAAAGTGTCTGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCA 335
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121  CCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCGGACTACTACGCCAAGGAGGTCA 180

Qy      336  CCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCC 395
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181  CCCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACAC 240

Qy      396  ACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGAAGCGGTGCCGGAACCTGTAT 455
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      241  ACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGT 300

Qy      456  TGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG 515
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      301  TGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG 360

Qy      516  AGCTATACCAGAAATACAGCAATGATTCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCC 575
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      361  AGCTGTACCAGAAATACAGCAACAAATCTTGGCGATACCTCAGCAACCGGCTGCTGGCAC 420

Qy      576  CCAGTGACTCACCGGAGTGGCTGTCTTTTGATGTACCCGGAGTTGTGCGGCAGTGGGCTGA 635
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      421  CCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTACCCGGAGTTGTGCGGCAGTGGTTGA 480

Qy      636  CCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAG 695
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      481  GCCGTGGAGGGGAAATTGAGGGCTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG 540
```

[illegible]

RESULT 13

BI818841

LOCUS	BI818841	925 bp	mRNA	linear	EST 04-OCT-2001
-------	----------	--------	------	--------	-----------------

DEFINITION 603037307F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5178433 5', mRNA sequence.

ACCESSION BI818841

VERSION BI818841.1 GI:15930391

KEYWORDS EST.

SOURCE	Homo sapiens (human)
--------	----------------------

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 925)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL      Unpublished

COMMENT            Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11444 row: n column: 02

High quality sequence stop: 874.

FEATURES	Location/Qualifiers
----------	---------------------

source

1. .925

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="IMAGE:5178433"
```

```
/lab_host="DH10B"
```

```
/clone lib="NIH MGC 115"
```

```
/note="Organ: pooled brain, lung, testis; Vector:
```

pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA

source anonymous pool of 6 male brains, age range 23-27; 1

male lung, age 27; and 1 male testis, age 69. Library

oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.8

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed b

Gruber (Invitrogen). Research Genetics tracking code

021. Note: this is a NIH\_MGC Library."  
BASE COUNT        170 a        316 c        278 g        161 t  
ORIGIN

Query Match                39.5%;    Score 537.6;    DB 12;    Length 925;  
Best Local Similarity    87.5%;    Pred. No. 3.3e-107;  
Matches 611;    Conservative    0;    Mismatches    84;    Indels        3;    Gaps        2;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  |||  |||  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     146 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 205

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    206 TGGTGCTGACGCCTGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 265

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    266 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 325

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 245
      |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    326 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCAGGCCGTGCTCG 385

Qy    246 CTCTTTACAACAGTACCCGCGACCGGTTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      |  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    386 CCCTGTACAACAGCACCCGCGACCGGTTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 445

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACC 365
      |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    446 CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACG 505

Qy    366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    506 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 565

Qy    426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    566 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 625

Qy    486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 545
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    626 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT 685

Qy    546 GGCGCTACCTCAGCAACCGGCTGC-TGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTT 604
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    686 GGCGATACCTCAGCAACCGGCTGCTTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTT 745

Qy    605 GATGTCAACCGAGTTGTGCGGCAGTGGCTGACCCGAGAG--GAGGCTATAGAGGGTTTTC 662
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    746 GATGTCAACCGAGTTGTGCGGCAGTGGTTGAGCCGTGGACGGGGAACATTGAGGGCTTTC 805

Qy    663 GCCTCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAAC 700
      |||  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    806 GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACC 843
```

RESULT 14  
AL530080/c

LOCUS AL530080 841 bp mRNA linear EST 23-MAY-2003

DEFINITION AL530080 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens  
cDNA clone CS0DD009YM06 3-PRIME, mRNA sequence.

ACCESSION AL530080

VERSION AL530080.2 GI:31067915

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 841)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:12793573.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

[cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r). Contact :

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD009BG03NP1.

FEATURES Location/Qualifiers

source

1..841

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DD009YM06"

/tissue\_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 150 a 240 c 283 g 153 t 15 others

ORIGIN

Query Match 38.6%; Score 525; DB 9; Length 841;

Best Local Similarity 80.3%; Pred. No. 1.8e-104;

Matches 689; Conservative 8; Mismatches 104; Indels 57; Gaps 5;

Qy 438 CGGTGCCCGGAACCTGTATTGCTCT-CTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTA 496  
|||| || |:|| | |||:||| | ||||| ||||| ||||| |||||

Db 841 CGGTACCTGMAACCTTGTGTCTCCCCGGGCAGAGCTGCTTCTGCTAAGGCTCAAGTTA 782

Qy 497 AAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTC 556  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 781 AAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATWCCTGGCGGATACCTC 722

Qy 557 AGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTTGATGTACCGGA 616



```

      |||||:| ||||| ||||| :||| || ||||| || ||||| |||||:|
Db      721 AGCAACCSGCSGCTGGCACCCAGCSACTCGCCAGAGTGGTTTCTTTGTTGTCACCGKA 662
Qy      617 GTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCAC 676
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      661 GTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTAAGG--TTTCGCCTTAGCGCCNAC 604
Qy      677 TGTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGC 736
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      603 GGCTCCTGTAAACAGCAGGGATAACACACTGCAAGTGAACATCAACGGGTTAACTACCGGC 544
Qy      737 CGCCGGGGTGACCTGGCCACCATTTCACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCC 796
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      543 CGCCGAGGTGACCTGGCCACCATTCTTGGCATGAACCGGCCCTTCCTGCTTCTCATGGCC 484
Qy      797 ACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATAACC 856
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      483 ACCCCGCTGGAGAGGGGCCAGCTTCTGCAAAGCTCCCGGCACCGCCGA----- 436
Qy      857 AACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGC 916
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      435 -----GCCCTGGACACCAACTATTGC 415
Qy      917 TTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGAC 976
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      414 TTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTAACCTCCGCAAGGAC 355
Qy      977 CTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCC 1036
      || ||||| ||||| || || ||||| |||||: ||||| ||||| |||||
Db      354 CTCGGCTGGGAGTGGATCCACGAGCCCAAGGGCTWCCATGCCAACTTCTGCCTCGGGCCC 295
Qy      1037 TGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGCTCTGTACAACCAG 1096
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      294 TGCCCTACATTTGGAGCCTGGACACGCAGGACAGCAAGGTCTTGCCCTGTACAACCAG 235
Qy      1097 CACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCC 1156
      || ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      234 CATAACCCGG--GCCTCGGCGGCGCCGTGC-GCGTGCCGCAGGCGCTGGAGCCGCTGCCC 178
Qy      1157 ATCGTGTAATACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGT 1216
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      177 TTCGTGTAATACTACGTGGGCGCGCAAGCCCAAGGTGGGGCGGCTGTCCAACATGGTTCGTGCGC 118
Qy      1217 TCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGG 1276
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      117 TCCTGCAAGGCGCGCTGGGGTCCCCGCCCCGCCCCGCCCCGCGCAGGCCCCGGCCCC 58
Qy      1277 CCCCACCCCCGCCCCGCT 1294
      |||| |||| |||: ||
Db      57 CCCCACCCCCGCCCCGCT 40

```

RESULT 15

BI084718/c

LOCUS

BI084718

956 bp

mRNA

linear

EST 20-JUN-2001

DEFINITION

602869722T1 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:5014421 3',  
mRNA sequence.

```

ACCESSION      BI084718
VERSION        BI084718.1  GI:14503048
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 956)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: ATCC
                cDNA Library Preparation: CLONETECH Laboratories, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLC1821 row: 1 column: 06
                High quality sequence start: 3
                High quality sequence stop: 793.
FEATURES       Location/Qualifiers
source         1..956
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5014421"
                /tissue_type="epidermoid carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_102"
                /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
                Site_2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed
                by Ling Hong in the laboratory of Gerald M. Rubin
                (University of California, Berkeley) using ZAP-cDNA
                synthesis kit (Stratagene) and Superscript II RT (Life
                Technologies). Note: this is a NIH_MGC Library."
BASE COUNT     172 a      273 c      312 g      199 t
ORIGIN
Query Match    38.5%; Score 524.6; DB 12; Length 956;
Best Local Similarity 77.0%; Pred. No. 2.3e-104;
Matches 757; Conservative 0; Mismatches 159; Indels 67; Gaps 7;

Qy      329 GAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATCTATGATAAATTCAAGGGC 388
      ||||| | ||| ||||||||| | ||| |
Db      955 GAGGTTAACCGGTGCTAAATGGTGGAACCCACCACGAATCTATGACCAGGTTTCAGCAGA 896

Qy      389 ACCCCCCACAGCTTATATATGCTGTTCAACACGTC--GGAGCTCCGGGAAGCGGTGCCGG 446
      | | | | || ||||| | ||| | | ||| ||| ||| ||
Db      895 GTACACACAGCCATAAATATGGTCTTCACACATTTCAGGAGCTTCCGAGAACCGGTACCCT 836

Qy      447 AACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTG-----CTGAGGCTCAAGTTAAAAGT 501
      | | | | || | | | | | | | | | | | | | | | |

```

Db 835 GAACCGTGTGGCTCTTCCCGGGGAAGAGCTGCCGTCTGCCTGAGGCTCAAGTTAAAAGT 776  
 Qy 502 -GGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCA 560  
 ||| ||||| ||||| ||| | | ||||| ||||| ||||| |||||  
 Db 775 GGGACCAGCACGTGGAGCTGTAAACAGGAATACAGCACAATTCTGGCGATACTCAGCA 716  
 Qy 561 ACCGGCTGCTGGCCCCCAGTGA--CTCACCAGGAGTGGCTGTCCTTTGATGTCACCGGAGT 618  
 ||||| ||||| ||||| ||| || ||||| | || ||||| ||||| |||||  
 Db 715 ACCGGCTGGTGGCACCCAGCGAACTCGCCAGAGTGGTTATCTTTGATGTCACCGGAGT 656  
 Qy 619 TGTGCGGCAGTGGCTGACCCGAG-AGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCACT 677  
 ||||| ||||| ||||| ||| ||| || || || ||||| ||||| ||||| |||||  
 Db 655 TGTGCGGCAGTGGTTGAGCCGTGGAAGGGGAAATTGAGGGCTTTCGCCTTAGCGCCCACT 596  
 Qy 678 GTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCC 737  
 | ||||| ||||| ||||| ||||| || ||||| || ||||| ||||| |||||  
 Db 595 GCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCC 536  
 Qy 738 GCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCA 797  
 |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 535 GCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCA 476  
 Qy 798 CCCCCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCA 857  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 475 CCCCCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC---- 420  
 Qy 858 ACAGCTACCCATACGACGTGCCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGCT 917  
 ||||| ||||| |||||  
 Db 419 -----ACCAACTATTGCT 407  
 Qy 918 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 977  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 406 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC 347  
 Qy 978 TGGGCTGGAAGTGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1037  
 | ||||| ||||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 346 TCGGCTGGAAGTGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCT 287  
 Qy 1038 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1097  
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 286 GCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 227  
 Qy 1098 ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1157  
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 226 ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA 167  
 Qy 1158 TCGTGTA TACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1217  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 166 TCGTGTA TACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 107  
 Qy 1218 CCTGCAAGTG CAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC 1277  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 106 CCTGCAAGTG CAGCTGAGGTCCCGCCCCGCC-----CCGCCCCGCCCGGCAGGCCCGGC 52  
 Qy 1278 CCCACCCCGCCCGCCTCACC GG 1300  
 ||||| ||||| ||||| |||||  
 Db 51 CCCACCCCGCCCGCCCCCGCTG 29

Search completed: October 28, 2003, 00:08:32  
Job time : 3384.39 secs

OM protein - protein search, using sw model

Run on: October 27, 2003, 19:11:58 ; Search time 33.9534 Seconds  
(without alignments)  
1902.657 Million cell updates/sec

Title: US-10-017-372E-39  
Perfect score: 2169  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	2038.5	94.0	390	23	AAE13596		Porcine transformi
2	1946.5	89.7	390	22	AAM39186		Human polypeptide
3	1939.5	89.4	390	7	AAP61468		PreTGF-beta gene p
4	1939.5	89.4	390	11	AAR04034		Sequence of pre-TG
5	1939.5	89.4	390	11	AAR05258		Human pre-transfor
6	1939.5	89.4	390	12	AAR13813		Human pro-TGF-beta
7	1939.5	89.4	390	16	AAR73596		Human TGF-beta 1 p
8	1939.5	89.4	390	17	AAR90827		Pre-transforming g
9	1936.5	89.3	390	23	AAU77101		Human transforming
10	1936.5	89.3	390	23	AAE16943		Human transforming
11	1936	89.3	391	24	ABB82780		TGFB1 Arg25Pro pol
12	1934.5	89.2	390	13	AAR20124		Sequence of simian
13	1930.5	89.0	390	15	AAR46227		Human pre-TGF-beta
14	1929	88.9	391	16	AAR83054		Transforming growt
15	1928.5	88.9	390	19	AAW78785		Human pre-transfor
16	1927.5	88.9	390	22	AAB84601		Nucleotide sequenc
17	1925	88.8	391	9	AAP81362		Human transforming
18	1924	88.7	434	11	AAR03743		Monkey transformin
19	1920.5	88.5	390	24	ABB82781		TGFB1 Arg25Pro pol
20	1903.5	87.8	386	11	AAR05663		Simian Transformin
21	1898	87.5	387	11	AAR05664		Simian Transformin
22	1887.5	87.0	390	11	AAR05492		Chimeric simian TG
23	1883.5	86.8	390	13	AAR27522		TGF-beta 1/beta 2
24	1864	85.9	389	13	AAR29657		TGF-beta 1. Homo
25	1859	85.7	453	22	AAM40972		Human polypeptide
26	1778.5	82.0	390	13	AAR20126		Sequence of hybrid
27	1777.5	82.0	390	11	AAR05749		Human TGF-Beta2 ex
28	1771.5	81.7	390	11	AAR05665		Human Transforming
29	1769.5	81.6	390	11	AAR05666		Hybrid transformin
30	1762	81.2	391	10	AAP91900		Sequence encoded b
31	1735	80.0	389	16	AAR79921		Simian-human hybri
32	1320	60.9	278	15	AAR53090		Polypeptide cross-
33	1315	60.6	278	12	AAR12541		Latency associated
34	1282.5	59.1	458	23	ABG31507		LAP-mIFNB construc
35	1282.5	59.1	463	23	ABG31510		LAP-huIFNB constru
36	1189	54.8	290	22	ABG06792		Novel human diagno
37	1155	53.3	450	23	ABG31508		mIFNB-LAP construc
38	1138	52.5	448	23	ABG31509		huIFNB-LAP construc
39	944	43.5	227	22	ABG20234		Novel human diagno
40	901	41.5	236	22	ABG20233		Novel human diagno
41	887.5	40.9	382	21	AAB08338		Amino acid sequenc
42	887.5	40.9	382	23	AAU77105		Frog transforming
43	849.5	39.2	456	19	AAW78786		Pig transforming g
44	845.5	39.0	412	16	AAR73598		Human TGF-beta 3 p
45	843	38.9	456	15	AAR46228		Pig TGF-beta-3. S

# ALIGNMENTS

RESULT 1

AAE13596

ID AAE13596 standard; Protein; 390 AA.

XX

AC AAE13596;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-beta1) mutant.

XX

KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;  
KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;  
KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; mutant;  
KW mutein.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT Misc-difference 223

FT /note= "Wild type Cys substituted with Ser"

FT Misc-difference 225

FT /note= "Wild type Cys substituted with Ser"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR N-PSDB; AAD22696.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel  
PT disease in humans, comprises vector containing transforming growth  
PT factor-beta under the control of inducible promoter -

XX

PS Example 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a  
CC gene encoding a regulatory transcription factor under the control of a  
CC promoter encoding a transforming growth factor-beta (TGF-beta). The  
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2  
CC or TGF-beta3, its variants or homologues, by transfecting a cell which  
CC is part of a host suspected of having an autoimmune disease, especially  
CC inflammatory bowel disease (IBD), under conditions such that the  
CC polypeptide encoded by the nucleic acid sequence in the vector is  
CC expressed. The vector is delivered using a delivery system. The delivery  
CC of the vector results in substantial elimination of symptoms of the  
CC autoimmune disease and increased production of IL-10 by the host. The  
CC composition is useful for treating various diseases with an autoimmune

CC component such as multiple sclerosis, rheumatoid arthritis, systemic  
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis  
CC and psoriasis, and also for assaying the expression of a gene in a cell.  
CC The vector is further useful for screening of the effect of test  
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.  
CC The present sequence is porcine TGF-beta1 mutant.

XX

SQ Sequence 390 AA;

Query Match 94.0%; Score 2038.5; DB 23; Length 390;  
Best Local Similarity 95.3%; Pred. No. 1.7e-170;  
Matches 388; Conservative 0; Mismatches 2; Indels 17; Gaps 1;

```
Qy      1 MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60
      |||
Db      1 MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||
Db     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKVKVEQHVELYQKYSNDSWR 180
      |||
Db    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKVKVEQHVELYQKYSNDSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      |||
Db    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      |||
Db    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTN 283

Qy    301 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      |||
Db    284 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      |||
Db    344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

## RESULT 2

AAM39186

ID AAM39186 standard; Protein; 390 AA.

XX

AC AAM39186;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2331.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;



KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI58342.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 4; SEQ ID NO 2331; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 SQ Sequence 390 AA;

Query Match 89.7%; Score 1946.5; DB 22; Length 390;  
 Best Local Similarity 90.4%; Pred. No. 2.1e-162;  
 Matches 368; Conservative 10; Mismatches 12; Indels 17; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Db	1	MPPSGLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTVGRQWLTRREAIEGFRLSAHCSCDSKDNLTlhVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTVGRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASIALDTN	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
Qy	361	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	344	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

# RESULT 3

AAP61468

ID AAP61468 standard; Protein; 390 AA.

XX

AC AAP61468;

XX

DT 31-OCT-2002 (updated)

DT 28-OCT-1991 (first entry)

XX

DE PreTGF-beta gene product.

XX

KW Transforming growth factor beta; cancer; wound healing.

XX

OS Unidentified.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Protein	279..390
----	---------	----------

XX

PN EP200341-A.

XX

PD 10-DEC-1986.

XX

PF 21-MAR-1986; 86EP-0302112.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

XX

PA (GETH ) GENENTECH INC.

XX



```
XX DT 25-MAR-2003 (updated)
DT 31-OCT-2002 (updated)
DT 31-MAY-1989 (first entry)
XX
DE Sequence of pre-TGF-beta 1.
XX
KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth
KW inhibition.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 348..500
XX
PN WO8912101-A.
XX
PD 14-DEC-1989.
XX
PF 08-JUN-1988; 88WO-US01945.
XX
PR 08-JUN-1988; 88WO-US01945.
XX
PA (GETH ) GENENTECH INC.
XX
PI Dernyck RMA, Goeddel DV;
XX
DR WPI; 1990-007474/01.
DR N-PSDB; AAQ02815.
XX
PT Nucleotide sequence encoding transforming growth factor beta-3 used as a
PT probe, or to produce TGF beta 3, for inhibiting growth of certain normal
PT and neoplastic cells, eg A549.
XX
PS Disclosure; Fig. 2; 6lpp; English.
XX
CC Sequence is an exon of transforming growth factor-beta 1 (pre-TGF-beta
CC 1) polypeptide and corresponds to AA's 288-338 of mature TGF-beta 1. The
CC nucleic acid encoding second subtype of TGF-beta (TGF-beta 3) is useful
CC as a probe or to produce TGF-beta 3 for inhibition of normal and
CC neoplastic cell growth.
CC (Updated on 31-OCT-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 390 AA;

Query Match 89.4%; Score 1939.5; DB 11; Length 390;
Best Local Similarity 90.2%; Pred. No. 8.5e-162;
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

QY 1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRRLA 60
Db | |||||
1 MPPSGLRLLPLLLPLLWLLVLTGPSPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRRLA 60
QY 61 SPSPSGQDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQI 120
| : |||||
```

Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAAEPEPEPEADYYAKEVTRVLMMVETHNEI	120
Qy	121	YDKFKGTPHSLYLNFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR       :    :  :	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN                                     :                         :           :	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWL SRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLL MATPLERAQHLHSSRHRRALD TNSYPYDV PDYAS LALDTN :	300
Db	241	TGRRGDLATIHGMNRPFLLL MATPLERAQHL QSSRHRR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW SLDTQYSKV LALY 	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW SLDTQYSKV LALY	343
Qy	361	NQHNP GASAAPCCVPQA LEPLPI VYYVG RKP KVEQL SNMIVRSCKCS	407
Db	344	NOHNPGASAAPCCVPOALEPLPI VYYVG RKP KVEQL SNMIVRSCKCS	390

## RESULT 5

AAR05258

ID AAR05258 standard; protein; 390 AA.

XX

AC AAR05258;

XX

DT 25-MAR-2003 (updated)

DT 05-AUG-1990 (first entry)

XX

DE Human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).

XX

KW Transforming growth factor-beta-1 (TGF-beta-1);

KW neoplastic cell line inhibition;

KW EGF-potentiated anchorage-independent growth.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Peptide	1..278
----	---------	--------

FT	Protein	279..2011
----	---------	-----------

FT	Domain	8..23
----	--------	-------

```
FT                               /note="hydrophobic domain"
```

FT Modified-site 82..84

FT	/note="potential N-glycosylation site "
----	---

FT	Modified-site	136..138
----	---------------	----------

FT /note="as above"

FT	Modified-site	176..178
----	---------------	----------

FT /note="as above"

FT	Cleavage-site	277..278
----	---------------	----------

FT /note="proteolytic cleavage site"

XX

PN US4886747-A.

XX

PD 12-DEC-1989.

XX  
PF 13-MAR-1987; 87US-0025423.  
XX  
PR 13-MAR-1987; 87US-0025423.  
PR 22-MAR-1985; 85US-0715142.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Derynck RMA, Goeddel DV;  
XX  
DR WPI; 1990-051338/07.  
DR N-PSDB; AAQ93301.  
XX  
PT Nucleic acid encoding transforming growth factor-beta -  
PT cloned into expression vectors for expression in eukaryotic host  
PT cells for therapeutic use  
XX  
PS Disclosure; Fig 1b; 28pp; English.  
XX  
CC Pre-TGF-beta-1 is a polypeptide of 390 amino acids. Post-translational  
CC cleavage of the precursor gives rise to the mature TGF-beta monomer.  
CC The sequence for human TGF-beta was determined by direct amino acid  
CC sequence analysis and by deduction from the TGF-beta cDNA. It is  
CC capable of inducing EGF-potentiated anchorage-independent growth of  
CC target cell lines, and/or growth inhibition of neoplastic cell lines. It  
CC can be used for treating wounds, eg burns or epidermal ulcers.  
CC (Updated on 25-MAR-2003 to correct PF field.)  
XX  
SQ Sequence 390 AA;

Qy 361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 6

AAR13813

ID AAR13813 standard; Protein; 390 AA.

XX

AC AAR13813;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1.

XX

KW Osteogenetic; tumoricidal.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Peptide	1..29
----	---------	-------

FT		/note= "signal peptide"
----	--	-------------------------

FT	Peptide	30..390
----	---------	---------

FT		/note= "pro-TGF-beta 1"
----	--	-------------------------

FT	Peptide	279..390
----	---------	----------

FT		/note= "TGF-beta 1"
----	--	---------------------

XX

PN JP03180192-A.

XX

PD 06-AUG-1991.

XX

PF 07-DEC-1989; 89JP-0318243.

XX

PR 07-DEC-1989; 89JP-0318243.

XX

PA (KIRI ) KIRIN BREWERY KK.

XX

DR WPI; 1991-271579/37.

DR N-PSDB; AAQ13392.

XX

PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by

PT preparing DNA chain contg. base sequence coding for human

PT pre:pro-TGF-beta 1, forming expression vector etc.

XX

PS Claim 1; Fig 1; 16pp; Japanese.

XX

CC The amino acid sequence codes for human prepro-TGF-beta 1 which

CC can be produced by recombinant methods, it has osteogenetic and

CC tumoricidal activity.

XX

SQ Sequence 390 AA;

Query Match 89.4%; Score 1939.5; DB 12; Length 390;

Best Local Similarity 90.2%; Pred. No. 8.5e-162;

Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIRGQILSKRLA	60
Db	1	MPPSGRLRLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKKRIEAIRGQILSKRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHCSCDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEI EGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATI HGMNRPFLLL MATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN	300
Db	241	TGRRGDLATI HGMNRPFLLL MATPLERAQHLQSSRHRR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
Qy	361	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	344	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

## RESULT 7

AAR73596

ID AAR73596 standard; Protein; 390 AA.

XX

AC AAR73596;

XX

DT 25-MAR-2003 (updated)

DT 20-DEC-1995 (first entry)

XX

DE Human TGF-beta 1 protein.

XX

KW Transforming growth factor-beta; Human TGF-beta protein; TGF-beta 1;

KW TGF-beta 3; osteogenic cell source; OCS; bone deficiency;

KW bone-inducing cofactor.

XX

OS Homo sapiens.

XX

PN US5409896-A.

XX

PD 25-APR-1995.

XX

PF 12-NOV-1993: 93US-0132405.

XX

PR 01-SEP-1989: 89US-0401906.

PR 12-NOV-1991; 91US-0790856.

PR 18-MAY-1993; 93US-0063841.

PR 12-NOV-1993; 93US-0132405.

XX





RESULT 8

AAR90827

ID AAR90827 standard; Protein; 390 AA.

XX

AC AAR90827;

XX

DT 25-MAR-2003 (updated)

DT 25-JAN-1980 (first entry)

XX

DE Pre-transforming growth factor beta 1.

XX

KW transforming growth factor beta 1; wound healing;

KW recombinant production.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Domain	8..23
----	--------	-------

FT		/note= "hydrophobic domain"
----	--	-----------------------------

FT	Modified-site	82..84
----	---------------	--------

FT		/note= "potential N-glycosylation site"
----	--	---

FT	Modified-site	136..138
----	---------------	----------

FT		/note= "potential N-glycosylation site"
----	--	---

FT	Modified-site	176..178
----	---------------	----------

FT		/note= "potential N-glycosylation site"
----	--	---

FT	Cleavage-site	277..279
----	---------------	----------

FT		/note= "trypsin-like peptidase cleavage site"
----	--	---

FT	Protein	279..390
----	---------	----------

FT		/label= mature_TGF_beta_1
----	--	---------------------------

XX

PN US5482851-A.

XX

PD 09-JAN-1996.

XX

PF 05-NOV-1993; 93US-0147364.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

PR 05-NOV-1993; 93US-0147364.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1996-076891/08.

DR N-PSDB; AAT15720.

XX

PT New recombinant human transforming growth factor-beta prods. - produced

PT using Chinese hamster ovary cells, for use in diagnostic applications

PT or in therapy

XX

PS Example 3; Fig 1A-C; 26pp; English.

XX

CC The pre-transforming growth factor (TGF) beta 1 protein is encoded  
CC by AAT15720. The mature TGF beta 1 monomer is cleaved from the  
CC precursor at the Arg-Arg dipeptide immediately preceding the mature  
CC TGF-beta 1 NH2-terminus. It does not contain a recognisable N-terminal  
CC signal peptide typical of most secreted proteins. The pre-TGF beta 1  
CC contains several pairs of basic residues which could undergo  
CC post-translational cleavage and give rise to separate polypeptide  
CC entities. The precursor contains 3 potential N-glycosylation sites, none  
CC of which are localised in the mature TGF beta 1. This is useful in  
CC purification of the mature protein. TGF beta 1 can be used in, e.g. wound  
CC healing.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 390 AA;

Query Match 89.4%; Score 1939.5; DB 17; Length 390;  
Best Local Similarity 90.2%; Pred. No. 8.5e-162;  
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

```
Qy      1 MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGRLRLPLLLPLLWLLVLTGPSPAAGLSTCKTIDMELVKRKRIEAI RGQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | | | : | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | | | | | | : | | | | | | | | | | : | | | | |
Db    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMRNPFLLLMATPLERAQHLQSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 9

AAU77101

ID AAU77101 standard; Protein; 390 AA.

XX

AC AAU77101;

XX

DT 05-JUN-2002 (first entry)

XX

DE Human transforming growth factor beta 1 (TFG-beta-1) polypeptide.

XX  
 KW Human; transforming growth factor beta; TGF-beta; insulin production;  
 KW type I diabetes mellitus; pancreatic cell outgrowth; wound healing;  
 KW pancreatic duct tissue; ischaemia; stroke; nervous system aging;  
 KW neurological condition; neurodegenerative disease; inflammation;  
 KW vasal injury; chemical injury; traumatic injury; tumour-induced injury;  
 KW amyotrophic lateral sclerosis; spinocerebellar degeneration;  
 KW immunological disease; multiple sclerosis; TGF-beta-1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200212336-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 09-FEB-2001; 2001WO-US04192.  
 XX  
 PR 09-AUG-2000; 2000US-0635368.  
 XX  
 PA (CURI-) CURIS INC.  
 XX  
 PI Wang M, Pang K;  
 XX  
 DR WPI; 2002-257468/30.  
 XX  
 PT Treating a subject with a disorder resulting from insufficient insulin  
 PT production, and inducing outgrowth of pancreatic cells, involves using  
 PT a transforming growth factor beta therapeutic -  
 XX  
 PS Disclosure; Fig 1; 77pp; English.  
 XX  
 CC The invention relates to treating a subject with a disorder resulting  
 CC from insufficient insulin production, involving contacting the subject  
 CC with a transforming growth factor beta (TGF-beta) therapeutic. TGF-beta  
 CC polypeptides can be used for treating a subject with a disorder resulting  
 CC from insufficient insulin production, e.g. type I diabetes mellitus, and  
 CC for inducing outgrowth of pancreatic cells associated with pancreatic  
 CC duct tissue within a subject. A composition comprising a TGF-beta protein  
 CC may be useful in wound healing and treatment of neurological conditions  
 CC derived from acute, subacute or chronic injury to the nervous system,  
 CC including traumatic injury, chemical injury, vasal injury and deficits  
 CC (such as ischaemia resulting from stroke), together with  
 CC infectious/inflammatory and tumour-induced injury, aging of the nervous  
 CC system including Alzheimer's disease, chronic neurodegenerative diseases  
 CC including Parkinson's disease, Huntington's chorea, amyotrophic lateral  
 CC sclerosis, spinocerebellar degenerations and chronic immunological  
 CC diseases of the nervous system or affecting the nervous system, including  
 CC multiple sclerosis. This sequence represents the human TGF-beta-1  
 CC protein.  
 XX  
 SQ Sequence 390 AA;

Query Match 89.3%; Score 1936.5; DB 23; Length 390;  
 Best Local Similarity 90.2%; Pred. No. 1.6e-161;  
 Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Db	1	MPPSGLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
QY	61	SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
QY	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
QY	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240
Db	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
QY	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSSRHRRALDTNSYPYDVPDYASLALDTN	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHR-----ALDTN	283
QY	301	YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
QY	361	NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	344	NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 10

AAE16943

ID AAE16943 standard; Protein; 390 AA.

XX

AC AAE16943;

XX

DT 18-APR-2002 (first entry)

XX

DE Human transforming growth factor-beta1 (TGF-beta1) protein.

XX

KW Human; transforming growth factor-beta1; TGF-beta1; osteoporosis;  
 KW latency associated peptide; LAP; integrin alphavbeta3; apoptosis;  
 KW immunomodulation; inflammatory disease; fibrotic disease; cancer;  
 KW diabetic retinopathy; chronic obstructive pulmonary disorder;  
 KW bone resorption; rheumatoid arthritis; psoriasis; restenosis;  
 KW atherosclerosis; liver fibrosis; asthma; cytostatic; osteopathic;  
 KW ophthalmological; antiarteriosclerotic; vasotropic.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal\_peptide

FT Region 30..278

FT /note= "LAP-beta1"

FT Domain 244..246

FT /note= "RGD motif"

FT Protein 279..390

FT /note= "Human mature TGF-beta1 protein"



```

Db      121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180
Qy      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
        |||
Db      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240
        |||
Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
        :|||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTN 283
        |||
Qy      301 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
        |||
Db      284 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343
        |||
Qy      361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
        |||
Db      344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
        |||

```

RESULT 11

ABB82780

ID ABB82780 standard; Protein; 391 AA.

XX

AC ABB82780;

XX

DT 18-MAR-2003 (first entry)

XX

DE TGFB1 Arg25Pro polymorphism G-allele protein sequence.

XX

KW Polymorphism; transforming growth factor beta 1; TGFB1; TGFbeta-1;

KW renal failure; nephrotropic; human; allele.

XX

OS Homo sapiens.

XX

PN WO200290585-A2.

XX

PD 14-NOV-2002.

XX

PF 08-MAY-2002; 2002WO-GB02066.

XX

PR 09-MAY-2001; 2001GB-0011277.

XX

PA (UYSH-) UNIV SHEFFIELD HALLAM.

XX

PI El-Nahas AM, Blakemore A, Khalil MS;

XX

DR WPI; 2003-120560/11.

DR N-PSDB; ABV75391.

XX

PT Determining an individual's susceptibility to the progression of renal

PT failure comprises detecting the presence of a genetic polymorphism

PT pattern in transforming growth factor beta 1 (TGFB1) gene in a sample

PT from the individual -

XX

PS Claim 51; Page 59-61; 62pp; English.

XX

CC The invention relates to determining an individual's susceptibility to





DE Sequence of simian transforming growth factor (TGF) beta-1.  
 XX  
 KW Hypertension therapy; hypotensive agent; blood pressure modulator.  
 XX  
 OS Monkey.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 8..21  
 FT Protein 279..390  
 XX  
 PN WO9119513-A.  
 XX  
 PD 26-DEC-1991.  
 XX  
 PF 20-JUN-1991; 91WO-US04449.  
 XX  
 PR 20-JUN-1990; 90US-0541221.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Oleson FB, Comerreski CR;  
 XX  
 DR WPI; 1992-024199/03.  
 DR N-PSDB; AAQ20289.  
 XX  
 PT Use of transforming growth factor (TGF)-beta and their  
 PT antagonists - for modulating blood pressure, for treating  
 PT hypertension and hypotension  
 XX  
 PS Disclosure; Fig 1; 42pp; English.  
 XX  
 CC A new method for treating hypertension comprises administering a  
 CC transforming growth factor (TGF)-beta to an individual at a dose  
 CC effective for lowering blood pressure; the TGF-beta may be e.g.  
 CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-  
 CC beta1 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-  
 CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2  
 CC complex.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 89.2%; Score 1934.5; DB 13; Length 390;  
 Best Local Similarity 89.9%; Pred. No. 2.3e-161;  
 Matches 366; Conservative 10; Mismatches 14; Indels 17; Gaps 1;

```

QY      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPSRPAAGLSTCKTIDMELVKRKRIETIRGQILSKLRLA 60

QY     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

QY    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

```

Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240
		:           :                       :	
Db	181	YLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSKDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN	300
		:	
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHQLQSSRHR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
Qy	361	NQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	344	NQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS	390

AAR46227

XX

XX

DT 09-JUL-1994 (first entry)

XX

XX

XX

XX

FT	Peptide	279..390
----	---------	----------

```
FT /label= Mat peptide
```

FT	Cleavage-site	279
----	---------------	-----

```
FT      /note= "TGF-beta-1 release site"
```

FT	Modified-site	82..84
----	---------------	--------

FT /label= N-glycosylation site

FT	Modified-site	136..138
----	---------------	----------

FT /label= N-glycosylation site

FT	Modified-site	176..178
----	---------------	----------

FT /label= N-glycosylation site

XX

XX

XX

XX

PR 13-MAR-1987; 87US-0025423.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992: 92US-0845893



RESULT 14

AAR83054

ID AAR83054 standard; Protein; 391 AA.

XX

AC AAR83054;

XX

DT 25-JUN-1996 (first entry)

XX

DE Transforming growth factor-beta 1.

XX

KW macrophage inducible nitric oxide synthase; iNOS; constitutive NOS;

KW interleukin-1-beta; transforming growth factor-beta; TGF-beta; IL1-beta;

KW nitric oxide production; hypotension; inflammation; septic shock;

KW treatment.

XX

OS Mammalian sp.

XX

FH Key Location/Qualifiers

FT Protein 279..391

FT /note= "represents the mature active TGF beta-1 mol."

XX

PN W09526745-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1994; 94WO-US03705.

XX

PR 05-APR-1994; 94WO-US03705.

XX

PA (HARD ) HARVARD COLLEGE.

XX

PI Lee M, Perrella MA;

XX

DR WPI; 1995-358443/46.

DR N-PSDB; AAT05876.

XX

PT Treatment of hypotension, esp. in septic shock - by administering

PT transforming growth factor-beta e.g. to inhibit inducible nitric

PT oxide synthase gene transcription

XX

PS Disclosure; Fig 17; 52pp; English.

XX

CC Transforming growth factor-beta 1 (TGF-beta 1) has been found to inhibit

CC inducible nitric oxide synthase (iNOS) gene transcription, esp. in

CC interleukin-1-beta (IL1-beta) stimulated rat smooth muscle cells, and at

CC a dose which does not inhibit constitutive NOS. TGF-beta 1 or 2

CC (AAR83055) or their active fragments (esp. derived from the

CC carboxy-terminal 112 amino acids), can be used in the treatment of

CC hypotension, such as that associated with severe inflammation or septic

CC shock.

XX

SQ Sequence 391 AA;

Query Match 88.9%; Score 1929; DB 16; Length 391;

Best Local Similarity 90.0%; Pred. No. 7.1e-161;

Matches 367; Conservative 10; Mismatches 13; Indels 18; Gaps 2;

Qy	1	MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLL-RLKLVKEQHVELYQKYSNDSW	179
		:     :   :	
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRRLKLVKEQHVELYQKYSNNSW	180
Qy	180	RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGF	239
		:                       :	
Db	181	RYLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGF	240
Qy	240	NSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDT	299
		:	
Db	241	TTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHR-----ALDT	283
Qy	300	NYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLAL	359
Db	284	NYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLAL	343
Qy	360	YNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	344	YNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS	391

# RESULT 15

AAW78785

ID AAW78785 standard; Protein; 390 AA.

XX

AC AAW78785;

XX

DT 25-MAR-2003 (updated)

DT 21-DEC-1998 (first entry)

XX

DE Human pre-transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; TGF-beta 1; human.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 8..23

FT /note= "hydrophobic domain"

FT Protein 279..390

FT /label= Mat\_protein

FT Modified-site 82..84

FT /note= "Asn is N-glycosylated"

FT Modified-site 136..138

FT /note= "Asn is N-glycosylated"

FT Modified-site 176..178

FT /note= "Asn is N-glycosylated"

FT Cleavage-site 277..278

```

FT      /note= "cleavage site for release of TGF-beta1"
XX
PN      US5801231-A.
XX
PD      01-SEP-1998.
XX
PF      30-MAY-1995;    95US-0454468.
XX
PR      13-MAR-1987;    87US-0025423.
PR      22-MAR-1985;    85US-0715142.
PR      04-AUG-1989;    89US-0389929.
PR      04-MAR-1992;    92US-0845893.
PR      05-NOV-1993;    93US-0147364.
PR      30-MAY-1995;    95US-0454468.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Derynck RMA,  Goeddel DV;
XX
DR      WPI; 1998-494840/42.
DR      N-PSDB; AAV52933.
XX
PT      DNA encoding transforming growth factor-beta precursor sequence -
PT      useful for analysis to perform manipulations to increase yield of
PT      recombinant production of the protein
XX
PS      Example 3; Fig 1B 1-3; 26pp; English.
XX
CC      This is the amino acid sequence of human transforming growth
CC      factor-beta 1 precursor (preTGF-beta 1). It was deduced from
CC      a preTGF-beta 1 cDNA sequence (see AAV52933). The invention relates
CC      to the recombinant production of TGF-beta. Biologically active
CC      TGF-beta is defined as being capable of inducing EGF-potentiated
CC      anchorage independent growth of target cell lines and/or growth
CC      inhibition of neoplastic cell lines. Nucleic acids encoding
CC      TGF-beta have been isolated and cloned into vectors which are
CC      replicated in bacteria and expressed in eukaryotic cells. TGF-beta
CC      recovered from transformed cells is used in known therapeutic
CC      applications.
CC      (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ      Sequence    390 AA;

Query Match          88.9%;  Score 1928.5;  DB 19;  Length 390;
Best Local Similarity 89.7%;  Pred. No. 7.9e-161;
Matches 365;  Conservative 10;  Mismatches 15;  Indels 17;  Gaps 1;

Qy      1  MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      |  |||||
Db      1  MPPSGLRLLPLLLPLLWLLVLTGPPAPGLSTCKTIDMEQVKRKRIEAIRGQILSKLRLA 60

Qy      61  SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||||:|||||
Db      61  SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      |||| : ||:|

```

```

Db      121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180
Qy      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      |||
Db      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240
Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      :|||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTN 283
Qy      301 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      |||
Db      284 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343
Qy      361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      |||
Db      344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

Search completed: October 28, 2003, 09:06:49  
Job time : 34.9534 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 09:09:54 ; Search time 24.6934 Seconds  
(without alignments)  
2760.110 Million cell updates/sec

Title: US-10-017-372E-39  
Perfect score: 2169  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*

```

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2050.5	94.5	390	11	US-09-214-592-26	Sequence 26, Appl
2	1949.5	89.9	390	11	US-09-214-592-29	Sequence 29, Appl
3	1946.5	89.7	390	15	US-10-087-268-2	Sequence 2, Appli
4	1939.5	89.4	390	15	US-10-087-268-5	Sequence 5, Appli
5	1936.5	89.3	390	12	US-10-276-947-1	Sequence 1, Appli
6	1934.5	89.2	390	11	US-09-214-592-33	Sequence 33, Appl
7	1929	88.9	391	11	US-09-214-592-17	Sequence 17, Appl
8	1925.5	88.8	390	10	US-09-756-283A-23	Sequence 23, Appl
9	1913.5	88.2	390	11	US-09-214-592-28	Sequence 28, Appl
10	1855.5	85.5	390	11	US-09-214-592-20	Sequence 20, Appl
11	1855.5	85.5	390	11	US-09-214-592-23	Sequence 23, Appl
12	1592.5	73.4	315	11	US-09-214-592-25	Sequence 25, Appl
13	1282.5	59.1	455	10	US-09-756-283A-20	Sequence 20, Appl
14	1159	53.4	447	10	US-09-756-283A-22	Sequence 22, Appl
15	1067	49.2	373	11	US-09-214-592-32	Sequence 32, Appl
16	888.5	41.0	412	11	US-09-214-592-31	Sequence 31, Appl
17	887.5	40.9	382	11	US-09-214-592-34	Sequence 34, Appl
18	887	40.9	383	10	US-09-756-283A-27	Sequence 27, Appl
19	851	39.2	409	11	US-09-214-592-27	Sequence 27, Appl
20	844.5	38.9	410	11	US-09-214-592-22	Sequence 22, Appl
21	844	38.9	414	11	US-09-214-592-21	Sequence 21, Appl
22	843.5	38.9	412	11	US-09-214-592-24	Sequence 24, Appl
23	841.5	38.8	412	11	US-09-214-592-19	Sequence 19, Appl
24	841.5	38.8	412	14	US-10-028-158-21	Sequence 21, Appl
25	837.5	38.6	412	10	US-09-756-283A-25	Sequence 25, Appl
26	829	38.2	414	10	US-09-756-283A-24	Sequence 24, Appl
27	829	38.2	414	11	US-09-214-592-18	Sequence 18, Appl
28	825	38.0	412	11	US-09-214-592-30	Sequence 30, Appl
29	812	37.4	304	10	US-09-756-283A-26	Sequence 26, Appl
30	753.5	34.7	139	14	US-10-002-278-8	Sequence 8, Appli
31	638	29.4	112	10	US-09-813-271B-2	Sequence 2, Appli
32	638	29.4	113	10	US-09-813-398-13	Sequence 13, Appl
33	638	29.4	114	10	US-09-813-459-22	Sequence 22, Appl
34	638	29.4	114	14	US-10-115-406-21	Sequence 21, Appl
35	638	29.4	114	15	US-10-154-333-23	Sequence 23, Appl



36	638	29.4	115	10	US-09-859-211-47	Sequence 47, Appl
37	638	29.4	115	10	US-09-880-708-25	Sequence 25, Appl
38	638	29.4	115	11	US-09-872-856-47	Sequence 47, Appl
39	638	29.4	115	15	US-10-335-483-29	Sequence 29, Appl
40	562	25.9	98	12	US-10-187-394-1	Sequence 1, Appli
41	542	25.0	116	14	US-10-115-406-24	Sequence 24, Appl
42	542	25.0	116	15	US-10-154-333-26	Sequence 26, Appl
43	535	24.7	112	10	US-09-813-271B-8	Sequence 8, Appli
44	504	23.2	112	10	US-09-813-271B-12	Sequence 12, Appl
45	504	23.2	114	14	US-10-115-406-25	Sequence 25, Appl

## ALIGNMENTS

## RESULT 1

```

US-09-214-592-26
; Sequence 26, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki,CMotoo
; APPLICANT: Shibata,CKenji
; APPLICANT: Sato,CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 26
; LENGTH: 390
; TYPE: PRT
; ORGANISM: porcine
US-09-214-592-26

```

Query Match 94.5%; Score 2050.5; DB 11; Length 390;  
Best Local Similarity 95.3%; Pred. No. 7.2e-183;  
Matches 388; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

Qy		1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA	60
Db		1 MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA	60
Qy		61 SPSPSGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMEESGNQI	120
		:	
Db		61 SPSPSGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMLESNGNI	120
Qy		121 YDKFKGTPHSLYMLFNTSELREAVEPVPVLSRAELRLLRLKLKVEQHVELYQKYNSDSWR	180
Db		121 YDKFKGTPHSLYMLFNTSELREAVEPVPVLSRAELRLLRLKLKVEQHVELYQKYNSDSWR	180
Qy		181 YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240
Db		181 YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240

```

Qy      241  SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
          |||||||||||||||||||||||||||||||||||||||||
Db      241  SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTN 283

Qy      301  YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
          |||||||||||||||||||||||||||||||||||||||||
Db      284  YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy      361  NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
          |||||||||||||||||||||||||||||||||||||||||
Db      344  NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 2

US-09-214-592-29

```

; Sequence 29, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki,CMotoo
; APPLICANT: Shibata,CKenji
; APPLICANT: Sato,CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 29
; LENGTH: 390
; TYPE: PRT
; ORGANISM: ovine
US-09-214-592-29

```

```

Query Match      89.9%; Score 1949.5; DB 11; Length 390;
Best Local Similarity 90.2%; Pred. No. 1.9e-173;
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

```

```

Qy      1  MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRKRIEAIHQILSKLRLA 60
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  MPPSGLRLLPLLLPLLWLLMLTPGRPVAGLSTCKTIDMELVKKRGIEAIHQILSKLRLA 60

Qy      61  SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  SPPSQGDVPPGGLPEAILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKI 120

Qy      121  YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
          | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121  YDKMKSSSHSIYMFNTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWR 180

Qy      181  YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181  YLSNRLLAPSDSPEWLSFDVTGVVRQWLTHREEIEGFRLSAHCSCDSKDNTLQVDINGFS 240

Qy      241  SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
          |||||||||||||||||||||||||||||||||||||||||

```

Db	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTN	283
Qy	301	YCFSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
Qy	361	NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKEQLSNMIVRSCKCS	407
Db	344	NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKEQLSNMIVRSCKCS	390

Query Match 89.7%; Score 1946.5; DB 15; Length 390;  
Best Local Similarity 90.4%; Pred. No. 3.7e-173;  
Matches 368; Conservative 10; Mismatches 12; Indels 17; Gaps 1;

```

Db      284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343
Qy      361 NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      344 NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 4

```

US-10-087-268-5
; Sequence 5, Application US/10087268
; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-10-087-268-5

```

```

Query Match      89.4%; Score 1939.5; DB 15; Length 390;
Best Local Similarity 90.2%; Pred. No. 1.7e-172;
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

```

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIKRGQILSKLRLA 60
      | ||||||||||||||||||| |||||||||||||||||||||||||||||
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPPAAGLSTCKTIDMELVKKRIEAIKRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||||:||||||||||||||| |||||||||||||||||||||:|:|
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      ||||| : ||:| ||||||||||||||||| |||||||||||||||||:|
Db    121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      |||||||||||||||||||:| |||||||||||||||:||| |:|
Db    181 YLSNRLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      :||||||||||||||||| ||||| |||||
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
      |||||||||||||||||||||||||||||||||||||||||||||||||||

```



# RESULT 6

US-09-214-592-33

; Sequence 33, Application US/09214592A  
 ; Publication No. US20030027218A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yamasaki,CMotoo  
 ; APPLICANT: Shibata,CKenji  
 ; APPLICANT: Sato,CYasufumi  
 ; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
 METHOD  
 ; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
 ; FILE REFERENCE: 11060  
 ; CURRENT APPLICATION NUMBER: US/09/214,592A  
 ; CURRENT FILING DATE: 1999-01-18  
 ; NUMBER OF SEQ ID NOS: 34  
 ; SOFTWARE:  
 ; SEQ ID NO 33  
 ; LENGTH: 390  
 ; TYPE: PRT  
 ; ORGANISM: simian  
 US-09-214-592-33

Query Match 89.2%; Score 1934.5; DB 11; Length 390;  
 Best Local Similarity 89.9%; Pred. No. 4.9e-172;  
 Matches 366; Conservative 10; Mismatches 14; Indels 17; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLT	PGRPAAGLSTCKTIDMELV	KRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLT	PSRPAAGLSTCKTIDMELV	KRKRIET	IRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYN	STRDRVAGESVEPEPEPEAD	YAKEVTRVLMVESGNQI		120
		:			:   :	
Db	61	SPPSQGEVPPGPLPEAVLALYN	STRDRVAGESAEPEPEPEAD	YAKEVTRVLMVETHNEI		120
Qy	121	YDKFKGTPHSLYMLFNTSELRE	AVPEPVLLSRAELRLLRLKL	KVEQHVELYQKYSNDSWR		180
		:     :			:	
Db	121	YDKFKQSTHSIYMFNTSELRE	AVPEPVLLSRAELRLLRLKL	KVEQHVELYQKYSNNSWR		180
Qy	181	YLSNRL LAPSDSPEWLSFDVT	GVVRQWLTRREAIEGFRLSA	HCSCDSKDNTLHVEINGFN		240
		:			:	
Db	181	YLSNRL LAPSNSPEWLSFDVT	GVVRQWLSRGGEIEGFRLSA	HCSCDSKDNTLQVDINGFT		240
Qy	241	SGRRGDLATIHGMNRPFLLLM	ATPLERAQHLHSSRHRRALD	TNSYPYDVPDYASLALDTN		300
		:				
Db	241	TGRRGDLATIHGMNRPFLLLM	ATPLERAQHLQSSRHRR-----	ALDTN		283
Qy	301	YCFSSTEKNCCVRQLYIDFRK	DLGWKWIHEPKGYHANFCLG	PCPYIWSLDTQYSKVLALY		360
Db	284	YCFSSTEKNCCVRQLYIDFRK	DLGWKWIHEPKGYHANFCLG	PCPYIWSLDTQYSKVLALY		343
Qy	361	NQHNP GASAAPCCVPQALEPL	PIVYVGRKPKVEQLSNMIVR	SCKCS		407
Db	344	NQHNP GASAAPCCVPQALEPL	PIVYVGRKPKVEQLSNMIVR	SCKCS		390

# RESULT 7



```
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-283A-23
```

```
Query Match          88.8%; Score 1925.5; DB 10; Length 390;
Best Local Similarity 89.7%; Pred. No. 3.4e-171;
Matches 365; Conservative 10; Mismatches 15; Indels 17; Gaps 1;
```

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAI R GQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPFPAAAGLSTCKTIDMELVKRKRIEAI R GQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | : : |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHHEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKFKQSTHSTYMFNISELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | : | | | | | | | | | : | | | : | | |
Db    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLL MATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLL MATPLERAQHLQSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

```
RESULT 9
US-09-214-592-28
; Sequence 28, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki, CMotoo
; APPLICANT: Shibata, CKenji
```



```

; APPLICANT: Sato,CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 28
; LENGTH: 390
; TYPE: PRT
; ORGANISM: canine
US-09-214-592-28

```

Qy	1	MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
		:	
Db	1	MPPSGLRLLPLLLPLLRLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLS	60
Qy	61	SPPSQGDVPVPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPPVPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVENTNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKEQHVELYQKYSNDSWR	180
		:     :     :	
Db	121	YEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRLKLKAEQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240
		:                       : :                       :       :	
Db	181	YLSNRLLAPSDTPEWL SFDVTGVVRQWL SHGGEVEGFRLSAHCSCDSKDNTLQVDINGFS	240
Qy	241	SGRRGDLATI HGMNRPFL LLMATPLERA QHLHSSRHRRALD TNSYPDV PDYAS LALDTN	300
Db	241	SSRRGDLATI HGMNRPFL LLMATPLERA QHLHSSRQR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
Qy	361	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	344	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

```

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 20
; LENGTH: 390
; TYPE: PRT
; ORGANISM: murine
US-09-214-592-20

```

```

Query Match          85.5%; Score 1855.5; DB 11; Length 390;
Best Local Similarity 85.5%; Pred. No. 1.2e-164;
Matches 348; Conservative 15; Mismatches 27; Indels 17; Gaps 1;

```

```

Qy      1 MAPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQLRKSSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDN TLHVEINGFN 240
      | | | | | | : | : | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLGNRLLTPDTPDWLSFDVTGVVRQWLNQGDGIQGFRFSAHCSCDSKDNKLHVEINGIS 240

Qy    241 SGRRGDLATI HGMNRPFLLL MATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTI HDMNRPFLLL MATPLERAQHLHSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNP GASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 11

US-09-214-592-23

```

; Sequence 23, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki,CMotoo
; APPLICANT: Shibata,Ckenji
; APPLICANT: Sato,Cyasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD

```

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 23  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: rat  
US-09-214-592-23

Query Match 85.5%; Score 1855.5; DB 11; Length 390;  
Best Local Similarity 85.5%; Pred. No. 1.2e-164;  
Matches 348; Conservative 14; Mismatches 28; Indels 17; Gaps 1;

```
QY      1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLPWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

QY     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESADPEPEPEPEADYYAKEVTRVLMVDRNNAI 120

QY    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | | : | | : | | | | | : | | | | | | | | | | | | | | | | |
Db    121 YDKTKDITHSIYMFNTSDIREAVPEPPLLSRAELRLQRFKSTVEQHVELYQKYSNNSWR 180

QY    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNLTlhVEINGFN 240
      | | | | | : | | : | | | | | | | | | | | : : | | | | | | | | | | :
Db    181 YLGNRLLTPDTPWLSFDVTGVVRQWLNQGDGIQGFRFSAHCSCDSKDNVLHVEINGIS 240

QY    241 SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTIHDNRNPFLLLMATPLERAQHLHSSRHRR-----ALDTN 283

QY    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

QY    361 NQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNPASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 12

US-09-214-592-25

; Sequence 25, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki, CMotoo  
; APPLICANT: Shibata, CKenji  
; APPLICANT: Sato, CYasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
METHOD  
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
; FILE REFERENCE: 11060

```

; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 25
; LENGTH: 315
; TYPE: PRT
; ORGANISM: bovine
US-09-214-592-25

```

Query Match 73.4%; Score 1592.5; DB 11; Length 315;  
Best Local Similarity 89.5%; Pred. No. 3.1e-140;  
Matches 297; Conservative 9; Mismatches 9; Indels 17; Gaps 1;

Qy	76	AVLALYNSTRDRVAGESVEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF	135
Db	1	A I L A L Y N S T R D R V A G E S A E T E P E P E A D Y Y A K E V T R V L M V E Y G N K I Y D K M K S S S H S I Y M F F	60
Qy	136	NTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEW	195
Db	61	NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLSNRLLAPSDSPEW	120
Qy	196	LSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFNSGRRGDLATIHGMNR	255
Db	121	LSFDVTGVVRQWLTRREEIEGFRLSAHCSCDSKDNTLQVDINGFSSGRRGDLATIHGMNR	180
Qy	256	PFLLLMATPLERAQHLHSSRHRRALDTSYPYDVPDYASLALDNTNYCFSSTEKNCCVRQL	315
Db	181	PFLLLMATPLERAQHLHSSRHR-----ALDNTNYCFSSTEKNCCVRQL	223
Qy	316	YIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVP	375
Db	224	YIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVP	283
Qy	376	QALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	284	QALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	315

```

RESULT 13
US-09-756-283A-20
; Sequence 20, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Artificial Sequence

```



Db 190 PLGLW---AGGGSAAALSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGEVPPGP 245

Qy 73 LPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLY 132  
 |||||:|:|:|

Db 246 LPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEIYDKFKQSTHSIY 305

Qy 133 MLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYSNDSWRYL SNRL LAPSD 191  
 |||||:|:|:|

Db 306 MFFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQHVELYQKYSNNSWRYL SNRL LAPSD 365

Qy 192 SPEWLSFDVTVGVVQWLTRREAIEGFRLSAHCSCDSKDN TLHVEINGFNSGRRGDLATIH 251  
 |||||:|:|:|

Db 366 SPEWLSFDVTVGVVQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFTTGRRGDLATIH 425

Qy 252 GMNRPFLLLMATPLERAQHLHS 273  
 |||||

Db 426 GMNRPFLLLMATPLERAQHLQS 447

RESULT 15

US-09-214-592-32

; Sequence 32, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki,CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato,CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 32

; LENGTH: 373

; TYPE: PRT

; ORGANISM: chicken

US-09-214-592-32

Query Match 49.2%; Score 1067; DB 11; Length 373;

Best Local Similarity 53.5%; Pred. No. 4.6e-91;

Matches 209; Conservative 52; Mismatches 98; Indels 32; Gaps 8;

Qy 30 LSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVA 89  
 |||||:|:|:|

Db 2 LSTCQRLDLEAAKKKRIEAVRGQILSKLRLTAPPPASETPPRPLPDDVRALYNSTQELLK 61

Qy 90 GES-VEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148  
 : : | | : : |||: | : | : : : : ||: : || | |

Db 62 QRARLRPPPDGPDEYWAKELRRI PMETTWDGAMEHWQPQSHSIFVFNVSRRRG-GRPT 120

Qy 149 LLSRAELRLLRLKLK-----VEQHVELYQKYSNDSWRYL SNRL LAPSDSPEWLSFDVTVGV 203  
 || |||||:| | ||:|||| | | |||| | : : |||||

Db 121 LLHRAELRMLRQKAAADSAGTEQRLELYQGYGNASWRYLHGRSVRATADDEWLSFDVTDA 180

```

Qy      204 VRQWLTRREAIEGFRLSAHCSCD---SKDNTLHVEINGFNSGRRGDLATIHGMNR--PFL 258
      | |||: | : || || |:      : : || | :|||: :| :| |::
Db      181 VHQLSGSELLGVFKLSVHCPCEMGPGHAEEMRISIEGFEQ-QRGDMQSIKKHRRVPYV 239

Qy      259 LLMATPLERAQHLHSSRRRRALDTNSYPYDVPDYASIALDTNYCF--SSTEKNCCVRQLY 316
      | || | ||| |||:| ||      |||:||| : ||||| |||
Db      240 LAMALPAERANELHSARRRR-----DLDTDYCFGPGTDEKNCCVRPLY 282

Qy      317 IDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQ 376
      ||||| ||||| |||||:||||| |||||:||||| ||||| |||||
Db      283 IDFRKDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNPASAAAPCCVPQ 342

Qy      377 ALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      |:||||:||||| :|||||||:|:|||||
Db      343 TLDPLPIIYYVGRNVRVEQLSNMIVRACKCS 373

```

Search completed: October 28, 2003, 09:28:02  
Job time : 24.6934 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:37:16 ; Search time 13.4491 Seconds  
(without alignments)  
2910.285 Million cell updates/sec

Title: US-10-017-372E-39  
Perfect score: 2169  
Sequence: 1 MAPSGLRLLPLLLPLWLLV.....GRKPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	2053.5	94.7	390	2	A27512	transforming growth	
2	2022	93.2	391	2	S01413	transforming growth	
3	1949.5	89.9	390	2	I46463	transforming growth	
4	1936.5	89.3	390	1	WFHU2	transforming growth	
5	1934.5	89.2	390	2	A26960	transforming growth	
6	1913.5	88.2	390	2	JC4023	transforming growth	
7	1855.5	85.5	390	1	WFMS2	transforming growth	
8	1855.5	85.5	390	2	S10219	transforming growth	
9	1592.5	73.4	315	2	A40057	transforming growth	
10	1067	49.2	373	2	A41918	transforming growth	
11	888.5	41.0	412	2	A34939	transforming growth	
12	887.5	40.9	382	2	B61036	transforming growth	
13	851	39.2	409	2	S01825	transforming growth	
14	844.5	38.9	410	2	A41397	transforming growth	
15	844	38.9	414	1	WFMSE2	transforming growth	
16	843.5	38.9	410	2	A55706	transforming growth	
17	841.5	38.8	412	2	A36169	transforming growth	
18	829	38.2	414	1	WFMKB2	transforming growth	
19	829	38.2	414	2	A31249	transforming growth	
20	828.5	38.2	413	1	WFXLB2	transforming growth	
21	825	38.0	412	2	A39489	transforming growth	
22	813	37.5	442	2	B31249	transforming growth	
23	691.5	31.9	130	2	I48196	transforming growth	
24	482	22.2	112	2	A61439	transforming growth	
25	275	12.7	425	2	I47072	inhibin beta-A chain	
26	270.5	12.5	424	1	WFPGBA	inhibin beta-A chain	
27	270	12.4	425	1	S50898	inhibin beta-A chain	
28	268.5	12.4	424	1	S31440	inhibin beta-A chain	
29	264.5	12.2	424	1	B40905	inhibin beta-A chain	
30	262.5	12.1	426	1	B24248	inhibin beta-A chain	
31	256.5	11.8	513	1	BMHU6	bone morphogenetic	
32	253.5	11.7	413	2	JC4862	activin beta-A chain	
33	251.5	11.6	398	2	JH0688	bone morphogenetic	
34	250	11.5	394	2	S45355	bone morphogenetic	
35	250	11.5	398	2	JH0687	bone morphogenetic	
36	249	11.5	396	1	BMHU2	bone morphogenetic	
37	246	11.3	510	2	A54798	Vg-1-related protein	
38	244	11.2	455	2	A43918	TGF-beta-related protein	
39	241.5	11.1	393	2	S37073	bone morphogenetic	
40	238	11.0	350	2	JC5241	activin beta E chain	
41	236	10.9	407	1	A40150	inhibin beta-B chain	
42	234.5	10.8	402	2	A45056	osteogenic protein	
43	232	10.7	411	2	B41398	inhibin beta-B chain	
44	229	10.6	393	2	I50103	activin beta B - z	
45	229	10.6	420	2	I49541	bone morphogenetic	

## ALIGNMENTS



A27512  
transforming growth factor beta-1 precursor - pig  
N;Alternate names: TGF-beta  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 05-Jun-1988 #sequence\_revision 05-Jun-1988 #text\_change 16-Jul-1999  
C;Accession: A27512; A26356; I46657  
R;Derynck, R.; Rhee, L.  
Nucleic Acids Res. 15, 3187, 1987  
A;Title: Sequence of the porcine transforming growth factor-beta precursor.  
A;Reference number: A27512; MUID:87174844; PMID:3470708  
A;Accession: A27512  
A;Molecule type: mRNA  
A;Residues: 1-390 <DER>  
R;Cheifetz, S.; Weatherbee, J.A.; Tsang, M.L.S.; Anderson, J.K.; Mole, J.E.;  
Lucas, R.; Massague, J.  
Cell 48, 409-415, 1987  
A;Title: The transforming growth factor-beta system, a complex pattern of cross-  
reactive ligands and receptors.  
A;Reference number: A90890; MUID:87102890; PMID:2879635  
A;Accession: A26356  
A;Molecule type: protein  
A;Residues: 279-322 <CHE>  
R;Kondaiah, P.; Van Obberghen-Schilling, E.; Ludwig, R.L.; Dhar, R.; Sporn,  
M.B.; Roberts, A.B.  
J. Biol. Chem. 263, 18313-18317, 1988  
A;Title: cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
Evidence for alternate splicing and polyadenylation.  
A;Reference number: I46657; MUID:89054010; PMID:2461367  
A;Accession: I46657  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-390 <KON>  
A;Cross-references: GB:M23703; NID:g755044; PIDN:AAA64616.1; PID:g755045  
C;Genetics:  
A;Gene: TGFB; TGF-beta-1  
C;Superfamily: inhibin  
C;Keywords: growth factor

Query Match 94.7%; Score 2053.5; DB 2; Length 390;  
Best Local Similarity 95.6%; Pred. No. 2.1e-158;  
Matches 389; Conservative 0; Mismatches 1; Indels 17; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMMVESGNQI	120
Db	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMMVESGNQI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQHVELY	QKYSNDSWR	180
Db	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQHVELY	QKYSNDSWR	180
Qy	181	YLSNRL LAPSDSPEWLSFDVTGV	VRQWLTRREAIEGFRLSAHCSCDSK	DNTHVEINGFN	240
Db	181	YLSNRL LAPSDSPEWLSFDVTGV	VRQWLTRREAIEGFRLSAHCSCDSK	DNTHVEINGFN	240

```

Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      |||
Db      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTN 283

Qy      301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      |||
Db      284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy      361 NQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
      |||
Db      344 NQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 2

S01413

transforming growth factor beta-1 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 24-Nov-1999

C;Accession: S01413

R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.

Nucleic Acids Res. 16, 8730, 1988

A;Title: Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1).

A;Reference number: S01413; MUID:88335639; PMID:3166520

A;Accession: S01413

A;Molecule type: DNA

A;Residues: 1-391 <JAK>

A;Cross-references: EMBL:X12373; NID:g63808; PIDN:CAA30933.1; PID:g63809

C;Superfamily: inhibin

C;Keywords: growth factor

```

Query Match          93.2%;  Score 2022;  DB 2;  Length 391;
Best Local Similarity 94.6%;  Pred. No. 7.6e-156;
Matches 386;  Conservative 0;  Mismatches 4;  Indels 18;  Gaps 2;

```

```

Qy      1  MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLA 60
      |||
Db      1  MPPSGPGLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLA 60

Qy      61  SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||
Db      61  SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120

Qy      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      |||
Db      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWG 180

Qy      181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEIN-GF 239
      |||
Db      181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINAGF 240

Qy      240 NSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDT 299
      |||
Db      241 NSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDT 283

Qy      300 NYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLAL 359

```

Db 284 NYCFSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLAL 343  
 Qy 360 YNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407  
 Db 344 YNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 391

### RESULT 3

I46463

transforming growth factor beta-1 - sheep

C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C;Date: 19-Dec-1997 #sequence revision 19-Dec-1997 #text change 24-Nov-1999

C;Accession: I46463; S45115

R; Woodall, C.J.; McLaren, L.J.; Watt, N.J.

Gene 150, 371-373, 1994

A;Title: Sequence and chromosomal localisation of the gene encoding ovine latent transforming growth factor-beta 1.

A;Reference number: I46463; MUID:95121932; PMID:7821809

A;Accession: I46463

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-390 <WOO>

A;Cross-references: EMBL:X76916; NID:q496648; PIDN:CAA54242.1; PID:q496649

A;Note: submitted to the EMBL Data Library, December 1993

C; Superfamily: inhibin

Query Match 89.9%; Score 1949.5; DB 2; Length 390;  
Best Local Similarity 90.2%; Pred. No. 5.6e-150;  
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

[illegible]

Db

344 NQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4

WFHU2

transforming growth factor beta-1 precursor [validated] - human

N;Alternate names: growth-inhibitory factor; TGF type 2; TGF-beta

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence\_revision 19-Oct-1995 #text\_change 08-Dec-2000

C;Accession: A27513; A01395; A22290; I59664; S53444

R;Derynck, R.; Rhee, L.; Chen, E.Y.; Van Tilburg, A.

Nucleic Acids Res. 15, 3188-3189, 1987

A;Title: Intron-exon structure of the human transforming growth factor-beta precursor gene.

A;Reference number: A27513; MUID:87174845; PMID:3470709

A;Accession: A27513

A;Molecule type: DNA

A;Residues: 1-390 <DER>

A;Cross-references: GB:X05839; GB:Y00112; NID:g37097; PIDN:CAA29283.1;

PID:g1212989

R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Eaton, D.H.; Bell, J.R.; Assoian, R.K.; Roberts, A.B.; Sporn, M.B.; Goeddel, D.V.

Nature 316, 701-705, 1985

A;Title: Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells.

A;Reference number: A01395; MUID:85296301; PMID:3861940

A;Accession: A01395

A;Molecule type: mRNA

A;Residues: 1-9,'P',11-24,'P',26-159,'R',160-390 <DE2>

A;Cross-references: GB:X02812; GB:J05114; NID:g37092; PIDN:CAA26580.1;

PID:g37093

A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide

R;Massague, J.; Like, B.

J. Biol. Chem. 260, 2636-2645, 1985

A;Title: Cellular receptors for type beta transforming growth factor. Ligand binding and affinity labeling in human and rodent cell lines.

A;Reference number: A22290; MUID:85131019; PMID:2982829

A;Accession: A22290

A;Molecule type: protein

A;Residues: 279-295,'XX',298-301 <MAS>

R;Urushizaki, Y.; Niitsu, Y.; Terui, T.; Koshida, Y.; Mahara, K.; Kohgo, Y.; Urushizaki, I.; Takahashi, Y.; Ito, H.

Tumor Res. 22, 41-55, 1987

A;Title: Cloning and expression of the gene for human transforming growth factor-beta in Escherichia coli.

A;Reference number: I59664

A;Accession: I59664

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 279-390 <RES>

A;Cross-references: GB:M38449; NID:g339557; PIDN:AAA36735.1; PID:g339558

R;Stam, K.; Stewart, A.A.; Qu, G.Y.; Iwata, K.K.; Fenyoe, D.; Chait, B.T.; Marshak, D.R.; Haley, J.D.

Biochem. J. 305, 87-92, 1995

A;Title: Physical and biological characterization of a growth-inhibitory activity purified from the neuroepithelioma cell line A673.

A;Reference number: S53444; MUID:95126934; PMID:7826358  
A;Accession: S53444  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 279-297 <STA>  
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.  
C;Genetics:  
A;Gene: GDB:TGFB1; TGFB  
A;Cross-references: GDB:120729; OMIM:190180  
A;Map position: 19q13.2-19q13.2  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; homodimer; mitogen; transformation  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-278/Domain: propeptide #status predicted <PRO>  
F;244-246/Region: cell attachment (R-G-D) motif  
F;279-390/Product: transforming growth factor beta-1 #status experimental <MAT>  
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.3%; Score 1936.5; DB 1; Length 390;  
Best Local Similarity 90.2%; Pred. No. 6.3e-149;  
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAI EGFRLSAHCSCDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNRLLAPSDSPEWLSFDVTVGVVRQWLSRGGEI EGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLL MATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLL MATPLERAQHLQSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 5  
A26960  
transforming growth factor beta-1 precursor - green monkey  
C;Species: Cercopithecus aethiops (green monkey, grivet)

C;Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 24-Nov-1999  
 C;Accession: A26960  
 R;Sharples, K.; Plowman, G.D.; Rose, T.M.; Twardzik, D.R.; Purchio, A.F.  
 DNA 6, 239-244, 1987  
 A;Title: Cloning and sequence analysis of simian transforming growth factor-beta  
 cDNA.  
 A;Reference number: A26960; MUID:87246074; PMID:3474130  
 A;Accession: A26960  
 A;Molecule type: mRNA  
 A;Residues: 1-390 <SHA>  
 A;Cross-references: GB:M16658; NID:g176552; PIDN:AAA35369.1; PID:g176553  
 C;Superfamily: inhibin  
 C;Keywords: growth factor  
 F;1-16/Domain: signal sequence #status predicted <SIG>  
 F;17-390/Product: transforming growth factor beta #status predicted <MAT>

Query Match 89.2%; Score 1934.5; DB 2; Length 390;  
 Best Local Similarity 89.9%; Pred. No. 9.1e-149;  
 Matches 366; Conservative 10; Mismatches 14; Indels 17; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPSRPAAGLSTCKTIDMELVKRKRIETIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | : | | | | | | | | | | | | | | | | | | | | | | : | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      | | | | | | | : | | | | | | | | | | | | | | | | | | | | | | : | | |
Db    181 YLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSKDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 6  
 JC4023  
 transforming growth factor beta-1 - dog  
 C;Species: Canis lupus familiaris (dog)  
 C;Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 24-Nov-1999  
 C;Accession: JC4023  
 R;Manning, A.M.; Auchampach, J.A.; Drong, R.F.; Slightom, J.L.  
 Gene 155, 307-308, 1995



A;Title: The murine transforming growth factor-beta precursor.  
A;Reference number: A01396; MUID:86168129; PMID:3007454  
A;Accession: A01396  
A;Molecule type: mRNA  
A;Residues: 1-390 <DER>  
A;Cross-references: GB:M13177; NID:g201952; PIDN:AAA40423.1; PID:g201953  
A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide  
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen; transformation  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-278/Domain: propeptide #status predicted <PRO>  
F;244-246/Region: cell attachment (R-G-D) motif  
F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>  
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.5%; Score 1855.5; DB 1; Length 390;  
Best Local Similarity 85.5%; Pred. No. 2.3e-142;  
Matches 348; Conservative 15; Mismatches 27; Indels 17; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | : | |
Db     61 SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | : | | | | | | | | | | | | | | | | | | | | | | | | | : | |
Db    121 YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQRLKSSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAHCSCDSKDNLTlhVEINGFN 240
      | | | | | | : | : | | | | | | | | | | | : : | : | | | | | | :
Db    181 YLGNRLLTPTDTPWLSFDVTVGVVRQWLNQGDGIQGFRFSAHCSCDSKDNKLHVEINGIS 240

Qy    241 SGRRGDLATIHGMNRPFLLLMTPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTIHDMNRPFLLLMTPLERAQHLHSSRHRR-----ALDTN 283

Qy    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

Qy    361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNP GASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 8  
S10219  
transforming growth factor beta-1 precursor - rat  
N;Alternate names: TGF type 2; TGF-beta





```

Db      241 PKRRGDLGTIHDNRPFLLLMATPLERAQHLHSSRHRR-----ALDTN 283
Qy      301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343
Qy      361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
        |||||:||||||||||||||||||||||||||||||||||
Db      344 NQHNP GASASPCCVQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 9

A40057

transforming growth factor beta-1 precursor - bovine (fragment)

N;Alternate names: beta-TGF; cartilage-inducing factor-A; EGF-dependent TGF or dEGF; MGF-b; milk growth factor b; TGF-type II

C;Species: Bos primigenius taurus (cattle)

C;Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 16-Jul-1999

C;Accession: A40057; A42320; A05284; A24322; B61439

R;Van Obberghen-Schilling, E.; Kondaiah, P.; Ludwig, R.L.; Sporn, M.B.; Baker, C.C.

Mol. Endocrinol. 1, 693-698, 1987

A;Title: Complementary deoxyribonucleic acid cloning of bovine transforming growth factor-beta1.

A;Reference number: A40057; MUID:91042552; PMID:3153459

A;Accession: A40057

A;Molecule type: mRNA

A;Residues: 1-315 <VAN>

A;Cross-references: GB:M36271; NID:g163747; PIDN:AAA30778.1; PID:g163748

R;Ogawa, Y.; Schmidt, D.K.; Dasch, J.R.; Chang, R.J.; Glaser, C.B.

J. Biol. Chem. 267, 2325-2328, 1992

A;Title: Purification and characterization of transforming growth factor-beta2.3 and -beta1.2 heterodimers from bovine bone.

A;Reference number: A42320; MUID:92129307; PMID:1733936

A;Accession: A42320

A;Molecule type: protein

A;Residues: 204-209,'X',211-217 <OGA>

R;Roberts, A.B.; Anzano, M.A.; Meyers, C.A.; Wideman, J.; Blacher, R.; Pan, Y.C.E.; Stein, S.; Lehrman, S.R.; Smith, J.M.; Lamb, L.C.; Sporn, M.B.

Biochemistry 22, 5692-5698, 1983

A;Title: Purification and properties of a type beta transforming growth factor from bovine kidney.

A;Reference number: A05284; MUID:84104793; PMID:6607069

A;Accession: A05284

A;Molecule type: protein

A;Residues: 204-218 <ROB>

R;Seyedin, S.M.; Thompson, A.Y.; Bentz, H.; Rosen, D.M.; McPherson, J.M.; Conti, A.; Siegel, N.R.; Galluppi, G.R.; Piez, K.A.

J. Biol. Chem. 261, 5693-5695, 1986

A;Title: Cartilage-inducing factor-A. Apparent identity to transforming growth factor-beta.

A;Reference number: A24322; MUID:86195954; PMID:3754555

A;Accession: A24322

A;Molecule type: protein

A;Residues: 204-233 <SEY>

R;Jin, Y.; Cox, D.A.; Knecht, R.; Raschdorf, F.; Cerletti, N.

J. Protein Chem. 10, 565-575, 1991

A;Title: Separation, purification, and sequence identification of TGF-beta1 and TGF-beta2 from bovine milk.  
A;Reference number: A61439; MUID:92189724; PMID:1799413  
A;Accession: B61439  
A;Molecule type: protein  
A;Residues: 204-209,'X',211-217,'XX',220-232 <JIN>  
C;Comment: This polypeptide is composed of two polypeptide chains cross-linked by disulfide bonds. It has been found in neoplastic and non-neoplastic tissues.  
C;Comment: Type II TGF does not bind to the EGF receptor and lacks intrinsic mitogenic activity, but in soft agar, it reacts synergistically with either type I TFG or EGF, and induces cell proliferation. Cells grown in monolayer do not respond in a similar manner to these growth factors, but morphologically do acquire a transformed phenotype.  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; heterodimer  
F;204-315/Product: transforming growth factor beta-1 #status experimental <MAT>  
F;7,61,101/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.4%; Score 1592.5; DB 2; Length 315;  
Best Local Similarity 89.5%; Pred. No. 3.2e-121;  
Matches 297; Conservative 9; Mismatches 9; Indels 17; Gaps 1;

```

Qy      76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135
      |:||||||||||||||| | ||||||||||||||||| ||:|||| | : ||:|||
Db      1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMF 60

Qy     136 NTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLNRLLAPSDSPEW 195
      |||||||||||||||:|||||||||||||:|||||||||
Db      61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLNRLLAPSDSPEW 120

Qy     196 LSFDTVGVVRQWLTRREAIEGFRLSAHCSCDSKDNTHLHVEINGFNSGRRGDLATIHGMNR 255
      ||||||||||||||||| ||||||||||||||||| |:||||:|||||||||
Db     121 LSFDTVGVVRQWLTRREEIEGFRLSAHCSCDSKDNTHLQVDINGFSSGRRGDLATIHGMNR 180

Qy     256 PFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTNYCFSSTEKNCCVRQL 315
      ||||||||||||||||| |||||||||||||||||
Db     181 PFLLLMATPLERAQHLHSSRHR-----ALDTNYCFSSTEKNCCVRQL 223

Qy     316 YIDFRKDLGKWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVP 375
      ||||||||||||||||| |||||||||||||||||
Db     224 YIDFRKDLGKWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVP 283

Qy     376 QALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      |||||||||||||||||
Db     284 QALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 315

```

RESULT 10  
A41918  
transforming growth factor beta-4 precursor - chicken (fragment)  
N;Alternate names: TGF-beta 4  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
C;Accession: A41918; A34941; S03110  
R;Burt, D.W.; Jakowlew, S.B.  
Mol. Endocrinol. 6, 989-992, 1992

A;Title: Correction: a new interpretation of a chicken transforming growth factor-beta 4 complementary DNA.  
A;Reference number: A41918; MUID:92357039; PMID:1353860  
A;Accession: A41918  
A;Molecule type: mRNA  
A;Residues: 1-373 <BUR>  
A;Cross-references: GB:M31160; GB:X08012; GB:S41706; NID:gl262437; PIDN:AAB05637.1; PID:gl262438  
A;Note: sequence extracted from NCBI backbone (NCBIN:110186, NCBIP:110187)  
A;Note: this report corrects and reinterprets the sequence from reference A34941 R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B. Mol. Endocrinol. 2, 1186-1195, 1988  
A;Title: Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid encoding transforming growth factor beta-4 from chicken embryo chondrocytes.  
A;Reference number: A34941; MUID:89112198; PMID:2464131  
A;Accession: A34941  
A;Molecule type: mRNA  
A;Residues: 'MDPMSIGPKSCGGSPWRPPGTAPWSIGSRRATASSSCSTSSRVRAEVGGRAL',122-209,'D',211-373 <JAK>  
A;Cross-references: EMBL:X08012  
A;Note: this sequence has been corrected in A41918  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor  
F;1/Domain: signal sequence (fragment) #status predicted <SIG>  
F;223-225/Region: cell attachment (R-G-D) motif  
F;260-373/Product: transforming growth factor beta-4 #status predicted <MAT>  
F;54,109,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.2%; Score 1067; DB 2; Length 373;  
Best Local Similarity 53.5%; Pred. No. 1.4e-78;  
Matches 209; Conservative 52; Mismatches 98; Indels 32; Gaps 8;

Qy	30	LSTCKTIDMELVKRKRIEAI	RQGILSKLRLASPPSQGDVPPG	PLPEAVLALYNSTRDRVA	89
		: :     :                     :	:         :                 : :		
Db	2	LSTCQRLDLEAAKKKRIEAV	RQGILSKLRLTAPPPASET	PPRPLPDDVRALYNSTQEL	61
Qy	90	GES-VEPEPEPEADYYAKE	VT	RVLMVESGNQIYDKFKGTPH	SLYMLFNTSELREAVPEPV
		: :     : :         :   : : : :     : :			
Db	62	QRARLRPPPDGPDEYWAKEL	RRI	PMETTWDGAMEHWQPQSHS	IFFVFNVSRARRG-GRPT
Qy	149	LLSRAELRLLRLK	----VEQHVELYQKYSNDS	WRYLSNRLLAPSDSPEWLS	SFDVTGV
		:	:	: :	
Db	121	LLHRAELRMLRQKAAADS	AGTEQRLELYQGYGNAS	WRYLHGRSVRATADDEWLS	SFDVTD
Qy	204	VRQWLTRREAIEGFRLSA	HCSCD---SKDNTLHVEING	FNSGRRGDLATI	HGMNR--PFL
		:   :	: :     :     :     :	:       :	
Db	181	VHQWLSGSELLGVFKLSV	HCPCEMGPGHAEEMRIS	IEGFEO-QRGDMOSI	AKKHRRVPYV
Qy	259	LLMATPLERAQHLHSSR	HRRALD	TNSYPYDVPDYASLALD	TNYCF--SSTEKNCCVRQ
		:		:       :	
Db	240	LAMALPAERANELHSARRR	-----DL	TDYCFGPGTDEKNCCVR	PLY
Qy	317	IDFRKDLGWKWIHEPKGY	HANFCLGPCPYIWSL	TQYSKVLALYNQHNPG	ASAAPCCVPQ
			:	:	
Db	283	IDFRKDLQWKWIHEPKGY	MANFCMGPCPYIWS	ADTQYTKVLALYNQHN	PGASAAPCCVPQ
			:	:	

QY 377 ALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407  
 |:||||:|||| :|||||:|:||||  
 Db 343 TLDPLPIIYYVGRNVRVEQLSNMVVRACKCS 373

RESULT 11

A34939

transforming growth factor beta-3 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 16-Jul-1999

C;Accession: A34939; S25850; S36125; S36124; I51181

R;Jakowlew, S.B.; Dillard, P.J.; Kondaiah, P.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 2, 747-755, 1988

A;Title: Complementary deoxyribonucleic acid cloning of a novel transforming growth factor-beta messenger ribonucleic acid from chick embryo chondrocytes.

A;Reference number: A34939; MUID:89096966; PMID:3211158

A;Accession: A34939

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-412 <JAK>

A;Cross-references: GB:M31154; NID:g212758; PIDN:AAA49089.1; PID:g212759

R;Burt, D.W.; Paton, I.R.; Dey, B.R.

J. Mol. Endocrinol. 7, 175-183, 1991

A;Title: Comparative analysis of human and chicken transforming growth factor-beta-2 and -beta-3 promoters.

A;Reference number: S25850; MUID:92134496; PMID:1840616

A;Accession: S25850

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-117 <BUR>

A;Cross-references: EMBL:X58127; NID:g63815; PIDN:CAA41128.1; PID:g63816

A;Accession: S36125

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 119-172 <BU2>

A;Cross-references: EMBL:X60055; NID:g396688; PIDN:CAA42653.1; PID:g396689

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

A;Accession: S36124

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 173-322,'ELPT',327-412 <BU3>

A;Cross-references: EMBL:X60091

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

R;Jakowlew, S.B.; Lechleider, R.; Geiser, A.G.; Kim, S.J.; Santa-Coloma, T.A.; Cubert, J.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 6, 1285-1298, 1992

A;Title: Identification and characterization of the chicken transforming growth factor-beta 3 promoter.

A;Reference number: I51181; MUID:93024487; PMID:1406706

A;Accession: I51181

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-117 <JA2>

A;Cross-references: GB:S46000; NID:g257172; PIDN:AAB23575.1; PID:g257173

C;Genetics:

A;Introns: 216/1; 252/1; 309/2; 360/3  
A;Note: list of introns may be incomplete  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-300/Domain: propeptide #status predicted <PRO>  
F;301-412/Product: transforming growth factor beta-3 #status predicted <MAT>  
F;74,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.0%; Score 888.5; DB 2; Length 412;  
Best Local Similarity 46.2%; Pred. No. 4.4e-64;  
Matches 198; Conservative 56; Mismatches 114; Indels 61; Gaps 13;

```

Qy      15 LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      9 LVLLSLLSFATVSLALSSCTTLDLEHIKKKRVEAI RQILSKLRLTSPPE--SVGPAHVP 66

Qy      75 EAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVE---SGNQIYDKF 124
      : ||||| : : | : | | | : ||||| : : | : : | : :
Db      67 YQILALYNSTRELL--EEMEEKEEESCSQENTESEYYAKEIHKFDMIQGLPEHNELGICP 124

Qy     125 KGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRL---KLKVEQHVELYQKYSND--- 177
      || : : | | | | | | | : ||| : || : | : | : | |
Db     125 KGVTSNVFR-FNVS---SAEKNSTNLFRAEFRVLRVPNPSSKRSEQRIELFQILRPDEHI 180

Qy     178 -SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC-----D 226
      |||| | : ||||| ||| ||| || : | || | :
Db     181 AKQRYLSGRNVQTRGSPEWLSFDVTDTVREWLLHRESNLGLEISIHCPCHTFQPNGDILE 240

Qy     227 SKDNTLHVEINGFNSG---RRGDLATI---HGMNRPFLLLMATPLERAQH--LHSSRHRR 278
      : | : | : | | | : : | : | | | : | | : |
Db     241 NLHEVLEIKFKGIDSEDDYGRGDLGRLKKQKDLHNPHLILMMLPPHRLESPTLGGQRKKR 300

Qy     279 ALDTNSYPYDVPDYASIALDNTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANF 338
      ||||| : | : |||| | ||||| : ||||| : ||||| |||
Db     301 -----ALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWVHEPKGYFANF 343

Qy     339 CLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSN 398
      | ||||| : | | : | | | | | | ||| : ||||| ||||| |||
Db     344 CSGPCPYLRSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSN 403

Qy     399 MIVRSCKCS 407
      | : | : |||||
Db     404 MVVKSCCKCS 412

```

# RESULT 12

B61036

transforming growth factor beta-5 precursor - African clawed frog

C;Species: *Xenopus laevis* (African clawed frog)

C;Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 16-Jul-1999

C;Accession: A34929; B61036

R;Kondaiah, P.; Sands, M.J.; Smith, J.M.; Fields, A.; Roberts, A.B.; Sporn, M.B.; Melton, D.A.

J. Biol. Chem. 265, 1089-1093, 1990

A;Title: Identification of a novel transforming growth factor-beta (TGF-beta5) mRNA in *Xenopus laevis*.

A;Reference number: A34929; MUID:90110090; PMID:2295601  
A;Accession: A34929  
A;Molecule type: mRNA  
A;Residues: 1-382 <KON>  
A;Cross-references: GB:J05180; NID:g214821; PIDN:AAA49968.1; PID:g214822  
R;Roberts, A.B.; Rosa, F.; Roche, N.S.; Coligan, J.E.; Garfield, M.; Rebbert, M.L.; Kondaiah, P.; Danielpour, D.; Kehrl, J.H.; Wahl, S.M.; Dawid, I.B.; Sporn, M.B.  
Growth Factors 2, 135-147, 1990  
A;Title: Isolation and characterization of TGF-beta2 and TGF-beta5 from medium conditioned by Xenopus XTC cells.  
A;Reference number: A61036; MUID:90253806; PMID:2340184  
A;Accession: B61036  
A;Molecule type: protein  
A;Residues: 271-276, 'X', 278-284, 'XX', 287-299 <ROB>  
C;Superfamily: inhibin  
C;Keywords: growth factor  
F;271-382/Product: transforming growth factor beta-5 #status experimental <MAT>

Query Match 40.9%; Score 887.5; DB 2; Length 382;  
Best Local Similarity 46.4%; Pred. No. 4.8e-64;  
Matches 192; Conservative 54; Mismatches 121; Indels 47; Gaps 11;

```

Qy      9 LPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDV 68
      : :| || |||| | |||| :||| |::| ||||| ||||| :| | |
Db      1 MEVLWMLLVLLVHLSSLAMSLSTCKAVDMEEVRKRRIEAIHQILSKLKLDKTPDV-DS 59

Qy     69 PPGPLPEAVLALYNSTRDRVAGESVEPE----PEPEADYYAKEVTRVLMVESGNQIYDK 123
      :| : ||||| : : : : | : ||||| :| | || : :|
Db     60 EKMTVPSEAIFLYNSTLEVIREKATREEEHVGHQNIQDYAKQVYR---FESITELED- 115

Qy    124 FKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKL--KVEQHVELYQKYSNDSW-- 179
      | || | :|| | || |||| : : : : :| :|| :| :
Db    116 -----HEFKFKFNASHVRENVGMNSLLHHAELRMYKKQTDKNMDQRMELFWKYQENGTT 170

Qy    180 -RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC---DSKDNTLHVE 235
      ||| : : | ||:|||| | :|| | | | | | | :|| : :
Db    171 SRYLESKYITPVTDDWMSFDVTKTVNEWLKRAEBENEQFGLQPACKCPTQAKD----ID 226

Qy    236 INGFNSGRRGDLATIHGM--NRPFLLLMATPLERAQHLHSSRRHRALDTNSYPYDVPDYA 293
      | || ||||| : : :| :| : : | || : ||| :|
Db    227 IEGF-PALRGDLASLSSKENTKPYLMTSMPAERIDTVTSSRKKR----- 270

Qy    294 SLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQY 353
      : ||| : |||| : ||| : ||| : ||| : ||| : ||| : |||
Db    271 --GVGQEYCFGNNGPNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPIYWSMDTQY 328

Qy    354 SKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      |||| : ||| : |||| : |||| : |||| : |||| : |||| : ||||
Db    329 SKVLSLYNQNNPGASISPCCVPDVLEPLPIIYYVGRTAKVEQLSNMVVRSCNCS 382

```

RESULT 13

S01825

transforming growth factor beta-3 precursor - pig

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999

C;Accession: S01825  
R;Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Mason, A.J.; Miller, D.A.; Coffey, R.J.; Moses, H.L.; Chen, E.Y.  
EMBO J. 7, 3737-3743, 1988  
A;Title: A new type of transforming growth factor-beta, TGF-beta3.  
A;Reference number: S01824; MUID:89091120; PMID:3208746  
A;Accession: S01825  
A;Molecule type: mRNA  
A;Residues: 1-409 <DER>  
A;Cross-references: EMBL:X14150; NID:g2127; PIDN:CAA32363.1; PID:g2128  
C;Superfamily: inhibin  
C;Keywords: growth factor  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;26-297/Domain: propeptide #status predicted <PRO>  
F;298-409/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match 39.2%; Score 851; DB 2; Length 409;  
Best Local Similarity 44.6%; Pred. No. 4.8e-61;  
Matches 190; Conservative 58; Mismatches 122; Indels 56; Gaps 13;

```
QY      15 LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      7 LVVLALLNFATVSLMSMSTCTTLD F DHIK RKRVEAIRQILSKLRLTSPDP SML--ANIP 64

QY      75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMV---ESGNQIYDKFKG 126
      || ||||| : | || : : | : ||||| : : | : | : | : ||
Db      65 TQVLDLYNSTRELLEEVHGERGDDCTQENT ESEYYAKEIYKFDMIQGLEEHNDLAVCPKG 124

QY      127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL----KLKVEQHVELYQKYSND----S 178
      : : || | : : | || | : || : : : | : | : | : | : |
Db      125 ITS KIFR-FNVSSVEK---NETNLFRAEFRLRMPNPSSKRSEQRIELFQILQPDEHIAK 180

QY      179 WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC-----DSK 228
      || : : | : ||||| || : || || : | : | || | : :
Db      181 QRYIDGKNLPTRGAAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENI 240

QY      229 DNTLHVEINGFNS---GRRGDLATIHGM--NRPFLLLMATPLERAQH--LHSSRHRRALD 281
      : : : | : | ||| : : | : || | : | : | : | : |
Db      241 QEVM EIKFKGVDS EDDPGRGDLGRLKKKKEHSPHLILMMIPDRLDN PGLGAQRKKR--- 297

QY      282 TNSYPYDVDPDYASLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLG 341
      ||||| : | : |||| | ||||| : ||||| : ||||| : |||||
Db      298 -----ALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQWVHEPKGYGANFCSG 343

QY      342 PCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIV 401
      ||| : | || : | || || || ||| : ||||| ||| : ||||| ||||| : |
Db      344 PCPYLRSADTT HSSVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGR TAKVEQLSNMVV 403

QY      402 RSCKCS 407
      : |||||
Db      404 KSCKCS 409
```

RESULT 14  
A41397  
transforming growth factor beta-3 precursor - mouse  
C;Species: Mus musculus (house mouse)



C;Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 16-Jul-1999  
 C;Accession: A41397; A61039; A61225  
 R;Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.  
 Mol. Endocrinol. 3, 1926-1934, 1989  
 A;Title: Complementary DNA cloning of the murine transforming growth factor-  
 beta3 (TGFbeta3) precursor and the comparative expression of TGFbeta3 and  
 TGFbeta1 messenger RNA in murine embryos and adult tissues.  
 A;Reference number: A41397; MUID:90190650; PMID:2628730  
 A;Accession: A41397  
 A;Molecule type: mRNA  
 A;Residues: 1-410 <MIL>  
 A;Cross-references: GB:M32745; NID:g201949; PIDN:AAA40422.1; PID:g201950  
 R;Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.  
 Growth Factors 3, 139-146, 1990  
 A;Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-  
 beta3.  
 A;Reference number: A61039; MUID:91000714; PMID:2206556  
 A;Accession: A61039  
 A;Molecule type: mRNA  
 A;Residues: 1-410 <DEN>  
 R;Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.  
 Cell Growth Differ. 2, 77-83, 1991  
 A;Title: Cell lineage specificity of expression of the murine transforming  
 growth factor beta-3 and transforming growth factor beta-1 genes.  
 A;Reference number: A61225; MUID:91299576; PMID:2069871  
 A;Accession: A61225  
 A;Status: translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 285-410 <WAT>  
 C;Superfamily: inhibin  
 C;Keywords: glycoprotein; growth factor; growth regulation  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-298/Domain: propeptide #status predicted <PRO>  
 F;259-261/Region: cell attachment (R-G-D) motif  
 F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>  
 F;72,133,140/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.9%; Score 844.5; DB 2; Length 410;  
 Best Local Similarity 44.2%; Pred. No. 1.6e-60;  
 Matches 188; Conservative 57; Mismatches 127; Indels 53; Gaps 12;

Qy	15	LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLASPPSQGDVPPGPLP	74
		: :	:          :  : : :	:
Db	7	LVLALLNLATISLSLSTCTTLD	FGHIKKKRVEAIRGQILSKLRLTSPPEPSVMT--HVP	64
Qy	75	EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---	SGNQIYDKFKG	126
			: :       :   :   :       : :   :   :	
Db	65	YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG		124
Qy	127	TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---	KLKVEQHVELYQKYSND----	S 178
		: :       : :         :   :   :   :		
Db	125	ITSKVFR-FNVSSVEK---NGTNLFRAEFRVLRVPNPSSKRTEQRIELFQILRPDEHIAK		180
Qy	179	WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC-----	DSK	228
		: :   :                 :   :	: :	
Db	181	QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV		240

QY 229 DNTLHVEINGFNS---GRRGDLATIHGM---NRPFLLLMATPLERAQHLHSSRHRRALDT 282  
 : : : | : : | | | : : | : | | | :  
 Db 241 HEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLI LMMI PPHRLDSPGQGSQRK--- 296

QY 283 NSYPYDVPDYASLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGP 342  
 | | | | | : | : | | | | | | : | | | | : | | | | |  
 Db 297 -----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWWHEPKGYANFCSGP 345

QY 343 CPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVR 402  
 | | : | | : | | | | | | : | | | | | | | : | :  
 Db 346 CPYLRADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVK 405

QY 403 SCKCS 407  
 | | | |  
 Db 406 SCKCS 410

# RESULT 15

WFMSB2

transforming growth factor beta-2 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 18-Jun-1999

C;Accession: A40148

R;Miller, D.A.; Lee, A.; Pelton, R.W.; Chen, E.Y.; Moses, H.L.; Derynck, R.

Mol. Endocrinol. 3, 1108-1114, 1989

A;Title: Murine transforming growth factor-beta2 cDNA sequence and expression in adult tissues and embryos.

A;Reference number: A40148; MUID:90014832; PMID:2797004

A;Accession: A40148

A;Molecule type: mRNA

A;Residues: 1-414 <MIL>

A;Cross-references: EMBL:X57413; NID:g54772; PIDN:CAA40672.1; PID:g54773

C;Comment: None of the three predicted glycosylation sites is in the mature protein.

C;Superfamily: inhibin

C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-302/Domain: propeptide #status predicted <PRO>

F;303-414/Product: transforming growth factor beta-2 #status predicted <MAT>

F;72,140,241/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.9%; Score 844; DB 1; Length 414;

Best Local Similarity 44.0%; Pred. No. 1.8e-60;

Matches 191; Conservative 61; Mismatches 120; Indels 62; Gaps 15;

QY 12 LLPLLWLLVLTPGRPAAGLSTCKTIDMELVKKRKRIEAI RQILSKLRLASPPSQGDVP-P 70  
 : | | | | | | | | : | : | | | | | | | | : | | | | | |  
 Db 5 VLSTFLLLHLVP--VALSLSTCSTLDMDQFM RKRRIEAI RQILSKLKLTSPP--DYEP 60

QY 71 GPLPEAVLALYNSTRD---RVAGESVEPEPE-PEADYYAKEVTRVLM---VESGNQIYD 122  
 : | | : : | | | | : : : | | | : | | | | : : | | | |  
 Db 61 DEVPPVISIYNSTRDLLQEKASRRAAACERERSEQEYYAKEVYKIDMPSHLPSENAIPP 120

QY 123 KFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----K 173  
 | | : : | : : : | : | | : | | | : | | : | | |  
 Db 121 TFY-RPYFRIVRFDVSTMEKNASN---LVKAEFRVFR LQNPKARVAEQRIELYQILKSKD 176

```

Qy      174 YSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC----- 225
      : : | : : : : | | | | | | | | : : | : : | | |
Db      177 LTSPTQRYIDSKVVKTRAEGEWLSFDVTDVQEWLHHKDRNLGFKISLHCPCTFVPSNN 236

Qy      226 ---DSKDNTLHVE---INGFNSGRRGDLATIHGMNR-----PFLLLMATPLERAQHLHS 273
      : | | | | : | : : | | | | : | | | | | | : |
Db      237 YIIPNKSEELERFAGIDGTSTYASGDQKTIKSTRKKTSGKTPHLLLMLLPSYRLESQQS 296

Qy      274 SRHRRALDTNSYPYDVPDYASLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKG 333
      | | : : | | | | | : : | | | : : | | | | | | | |
Db      297 SRRKKR-----ALDAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKG 340

Qy      334 YHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKV 393
      | : | | | | | | : | | | : | | : | | | | | | | | : | | : | | :
Db      341 YNANFCAGACPYLWSSDTQHTKVLSLYNTINPEASAPCCVSQDLEPLTILYYIGNTPKI 400

Qy      394 EQLSNMIVRSCKCS 407
      | | | | | | : | | | |
Db      401 EQLSNMIVKSKCS 414

```

Search completed: October 28, 2003, 09:09:50  
Job time : 14.4491 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:08:41 ; Search time 8.15764 Seconds  
(without alignments)  
2346.251 Million cell updates/sec

Title: US-10-017-372E-39  
Perfect score: 2169  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	2050.5	94.5	390	1	TGF1_PIG	P07200	sus scrofa
2	1949.5	89.9	390	1	TGF1_SHEEP	P50414	ovis aries
3	1936.5	89.3	390	1	TGF1_HUMAN	P01137	homo sapien
4	1934.5	89.2	390	1	TGF1_CERAE	P09533	cercopithec
5	1913.5	88.2	390	1	TGF1_CANFA	P54831	canis famil
6	1859.5	85.7	390	1	TGF1_HORSE	O19011	equus cabal
7	1855.5	85.5	390	1	TGF1_MOUSE	P04202	mus musculu
8	1855.5	85.5	390	1	TGF1_RAT	P17246	rattus norv
9	1846.5	85.1	390	1	TGF1_CAVPO	Q9zly6	cavia porce
10	1592.5	73.4	315	1	TGF1_BOVIN	P18341	bos taurus
11	1067	49.2	373	1	TGF1_CHICK	P09531	gallus gall
12	887.5	40.9	382	1	TGF1_XENLA	P16176	xenopus lae
13	869.5	40.1	412	1	TGF3_CHICK	P16047	gallus gall
14	851	39.2	409	1	TGF3_PIG	P15203	sus scrofa
15	844.5	38.9	410	1	TGF3_MOUSE	P17125	mus musculu
16	844	38.9	414	1	TGF2_MOUSE	P27090	mus musculu
17	843.5	38.9	412	1	TGF3_RAT	Q07258	rattus norv
18	841.5	38.8	412	1	TGF3_HUMAN	P10600	homo sapien
19	829	38.2	414	1	TGF2_HUMAN	P08112	homo sapien
20	828.5	38.2	413	1	TGF2_XENLA	P17247	xenopus lae
21	825	38.0	412	1	TGF2_CHICK	P30371	gallus gall
22	825	38.0	435	1	TGF2_PIG	P09858	sus scrofa
23	825	38.0	442	1	TGF2_RAT	Q07257	rattus norv
24	482	22.2	112	1	TGF2_BOVIN	P21214	bos taurus
25	323.5	14.9	375	1	GDF8_MELGA	O42221	meleagris g
26	321.5	14.8	375	1	GDF8_CHICK	O42220	gallus gall
27	312.5	14.4	375	1	GDF8_PIG	O18831	sus scrofa
28	311.5	14.4	375	1	GDF8_PAPHA	O18828	papio hamad
29	309.5	14.3	376	1	GDF8_RAT	O35312	rattus norv
30	308.5	14.2	375	1	GDF8_HUMAN	O14793	homo sapien
31	307.5	14.2	376	1	GDF8_MOUSE	O08689	mus musculu
32	305.5	14.1	375	1	GDF8_SHEEP	O18830	ovis aries
33	298	13.7	405	1	GDFB_MOUSE	Q9z1w4	mus musculu
34	298	13.7	407	1	GDFB_HUMAN	O95390	homo sapien
35	296.5	13.7	375	1	GDF8_BOVIN	O18836	bos taurus
36	280.5	12.9	374	1	GDF8_BRARE	O42222	brachydanio
37	276.5	12.7	345	1	GDFB_RAT	Q9z217	rattus norv
38	275	12.7	425	1	IHBA_SHEEP	P43032	ovis aries
39	271.5	12.5	426	1	IHBA_HORSE	P55102	equus cabal
40	270.5	12.5	424	1	IHBA_PIG	P03970	sus scrofa
41	270	12.4	425	1	IHBA_BOVIN	P07995	bos taurus
42	268.5	12.4	424	1	IHBA_MOUSE	Q04998	mus musculu
43	268.5	12.4	424	1	IHBA_RAT	P18331	rattus norv
44	265.5	12.2	424	1	IHBA_CHICK	P27092	gallus gall
45	263	12.1	424	1	BM10_HUMAN	O95393	homo sapien

## ALIGNMENTS

TGF1\_PIG

ID TGF1\_PIG STANDARD; PRT; 390 AA.

AC P07200; P08832;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transforming growth factor beta 1 precursor (TGF-beta 1).

GN TGFB1.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RX MEDLINE=87174844; PubMed=3470708;

RA Derynck R., Rhee L.;

RT "Sequence of the porcine transforming growth factor-beta precursor.";

RL Nucleic Acids Res. 15:3187-3187(1987).

RN [2]

RP SEQUENCE FROM N.A., AND VARIANT VAL-114.

RC STRAIN=Miniature swine;

RX MEDLINE=89054010; PubMed=2461367;

RA Kondaiah P., van Obberghen-Schilling E., Ludwig R.L., Dhar R.,

RA Sporn M.B., Roberts A.B.;

RT "cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.

RT Evidence for alternate splicing and polyadenylation.";

RL J. Biol. Chem. 263:18313-18317(1988).

RN [3]

RP SEQUENCE FROM N.A., AND VARIANT VAL-114.

RX MEDLINE=88335639; PubMed=3166520;

RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;

RT "Nucleotide sequence of chicken transforming growth factor-beta 1

RT (TGF-beta 1).";

RL Nucleic Acids Res. 16:8730-8730(1988).

RN [4]

RP SHOWS THAT REF.3 SEQUENCE IS FROM PIG.

RA Jakowlew S.B.;

RL Unpublished observations (MAR-1996).

RN [5]

RP SEQUENCE FROM N.A., AND VARIANT VAL-114.

RA Wimmers K., Chomdej S., Ponsuksili S., Schellander K.;

RT "Polymorphism in the porcine transforming growth factor beta 1

RT gene.";

RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.

RN [6]

RP SEQUENCE OF 279-322.

RX MEDLINE=87102890; PubMed=2879635;

RA Cheifetz S., Weatherbee J.A., Tsang M.L.S., Anderson J.K., Mole J.E.,

RA Lucas R., Massague J.;

RT "The transforming growth factor-beta system, a complex pattern of

RT cross-reactive ligands and receptors.";

RL Cell 48:409-415(1987).

CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL

CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL

CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM

CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE

CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -!- CAUTION: REF.3 SEQUENCE WHICH WAS SAID TO ORIGINATE FROM CHICKEN  
 CC WHITE LEGHORN, SEEMS (REF.4) TO ORIGINATE FROM PIG.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y00111; CAA68291.1; -.  
 DR EMBL; M23703; AAA64616.1; -.  
 DR EMBL; X12373; CAA30933.1; -.  
 DR EMBL; AF461808; AAL57902.1; -.  
 DR PIR; A27512; A27512.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal; Polymorphism.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 FT VARIANT 114 114 L -> V.  
 FT CONFLICT 6 7 LR -> PG (IN REF. 3).  
 FT CONFLICT 180 180 R -> G (IN REF. 3).  
 FT CONFLICT 237 237 N -> NA (IN REF. 3).  
 SQ SEQUENCE 390 AA; 44294 MW; A6E2C3659FC384E6 CRC64;

Query Match 94.5%; Score 2050.5; DB 1; Length 390;  
 Best Local Similarity 95.3%; Pred. No. 6.5e-157;  
 Matches 388; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQGILSKLR L A 60



CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X76916; CAA54242.1; -.

DR PIR: I46463; I46463.

DR InterPro; IPR002400; GF cysknot.

```
DB InterPro: IPR001839: TGFb
```

DR Pfam: PF00019: TGF-beta: 1.

DR Pfam; PF00688; TGFb propeptide; 1.

DR PRINTS; PR00438; GFCYSKNOT.

DR PRINTS; PR01423; TGFBETA.

DR ProDom; PD000357; TGFb; 1.

DR SMART; SM00204; TGFB; 1.

DR PROSITE; PS00250; TGF BETA 1; 1.

KW Growth factor; Mitogen; Glycoprotein; Signal.

FT      SIGNAL                      1            23            POTENTIAL.

FT PROPEP 24 278 POTENTIAL.

FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.

FT    DISULFID    285    294    BY SIMILARITY.

FT    DISULFID    293    356    BY SIMILARITY.

FT    DISULFID    322    387    BY SIMILARITY.

FT    DISULFID    326    389    BY SIMILARITY.

FT	DISULFID	355	355	INTERCHAIN (BY SIMILARITY).
----	----------	-----	-----	-----------------------------

FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
----	----------	----	----	-------------------------------------

FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
----	----------	-----	-----	-------------------------------------

FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . . ) (POTENTIAL)
----	----------	-----	-----	-------------------------------------

FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL)
----	------	-----	-----	----------------------------------

SO SEQUENCE 390 AA: 44291 MW: 1C247299484D0E57 CRC64:

Query Match 89.9%; Score 1949.5; DB 1; Length 390;

Best Local Similarity 90.2%; Pred. No. 8.1e-149;

Qy 1 MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Db 1 MPPSGLRLLPLLLPLLWLLMLTPGRPVAGLSTCKTIDMELVKRKGIEAIRGOILSKLRLA 60

Ov 61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNOI 120

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	5
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---



```

Db          61 SPSPQGDVPPGPLPEAILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKI 120
Qy          121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
          ||| | : ||:| | ||||| ||||| ||||| : ||||| ||||| ||||| : |||
Db          121 YDKMKSSSHSIYMFFNTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWR 180
Qy          181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHCSCDSKDNTLHVEINGFN 240
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : ||||| :
Db          181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTHREEI EGFRLSAHCSCDSKDNTLQVDINGFS 240
Qy          241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTN 283
Qy          301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343
Qy          361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 3

## TGF1\_HUMAN

```

ID   TGF1_HUMAN          STANDARD;          PRT;   390 AA.
AC   P01137; Q9UCG4;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFB1 OR TGFB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87174845; PubMed=3470709;
RA   Derynck R., Rhee L., Chen E.Y., van Tilburg A.;
RT   "Intron-exon structure of the human transforming growth factor-beta
RT   precursor gene.";
RL   Nucleic Acids Res. 15:3188-3189(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85296301; PubMed=3861940;
RA   Derynck R., Jarrett J.A., Chen E.Y., Eaton D.H., Bell J.R.,
RA   Assoian R.K., Roberts A.B., Sporn M.B., Goeddel D.V.;
RT   "Human transforming growth factor-beta complementary DNA sequence and
RT   expression in normal and transformed cells.";
RL   Nature 316:701-705(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Duodenum, and Eye;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 279-390 FROM N.A.  
 RC TISSUE=Carcinoma;  
 RA Urushizaki Y., Niitsu Y., Terui T., Koshida Y., Mahara K., Kohgo Y.,  
 RA Urushizaki I., Takahashi Y., Ito H.;  
 RT "Cloning and expression of the gene for human transforming growth  
 RT factor-beta in Escherichia coli.";  
 RL Tumor Res. 22:41-55(1987).  
 RN [5]  
 RP SEQUENCE OF 279-329.  
 RC TISSUE=Bladder carcinoma;  
 RX MEDLINE=93229900; PubMed=8471846;  
 RA Bourdrel L., Lin C.-H., Lauren S.L., Elmore R.H., Sugarman B.J.,  
 RA Hu S., Westcott K.R.;  
 RT "Recombinant human transforming growth factor-beta 1: expression by  
 RT Chinese hamster ovary cells, isolation, and characterization.";  
 RL Protein Expr. Purif. 4:130-140(1993).  
 RN [6]  
 RP SEQUENCE OF 279-301.  
 RX MEDLINE=85131019; PubMed=2982829;  
 RA Massague J., Like B.;  
 RT "Cellular receptors for type beta transforming growth factor. Ligand  
 RT binding and affinity labeling in human and rodent cell lines.";  
 RL J. Biol. Chem. 260:2636-2645(1985).  
 RN [7]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=93144319; PubMed=8424942;  
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: NMR signal assignments of the  
 RT recombinant protein expressed and isotopically enriched using Chinese  
 RT hamster ovary cells.";  
 RL Biochemistry 32:1152-1163(1993).  
 RN [8]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=93144320; PubMed=8424943;  
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,

RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: secondary structure as determined  
 RT by heteronuclear magnetic resonance spectroscopy.";   
 RL Biochemistry 32:1164-1171(1993).  
 RN [9]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=96266150; PubMed=8679613;  
 RA Hinck A.P., Archer S.J., Qian S.W., Roberts A.B., Sporn M.B.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: three-dimensional structure in  
 RT solution and comparison with the X-ray structure of transforming  
 RT growth factor beta 2.";   
 RL Biochemistry 35:8517-8534(1996).  
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,  
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY  
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE  
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A  
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X05839; CAA29283.1; -.  
 DR EMBL; X05840; CAA29283.1; JOINED.  
 DR EMBL; X05843; CAA29283.1; JOINED.  
 DR EMBL; X05844; CAA29283.1; JOINED.  
 DR EMBL; X05849; CAA29283.1; JOINED.  
 DR EMBL; X05850; CAA29283.1; JOINED.  
 DR EMBL; X02812; CAA26580.1; ALT\_SEQ.  
 DR EMBL; BC001180; AAH01180.1; -.  
 DR EMBL; BC000125; AAH00125.1; -.  
 DR EMBL; BC022242; AAH22242.1; -.  
 DR EMBL; M38449; AAA36735.1; -.  
 DR PIR; A27513; WFHU2.  
 DR PDB; 1KLA; 17-AUG-96.  
 DR PDB; 1KLC; 17-AUG-96.  
 DR PDB; 1KLD; 17-AUG-96.  
 DR Genew; HGNC:11766; TGFB1.  
 DR MIM; 190180; -.  
 DR GO; GO:0006916; P:anti-apoptosis; TAS.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.

```

DR PRINTS; PR01423; TGFbeta.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 278
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
FT DISULFID 285 294
FT DISULFID 293 356
FT DISULFID 322 387
FT DISULFID 326 389
FT DISULFID 355 355 INTERCHAIN.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
FT CONFLICT 10 10 L -> P (IN REF. 2).
FT CONFLICT 159 159 R -> RR (IN REF. 2).
FT STRAND 281 281
FT TURN 282 287
FT STRAND 294 296
FT STRAND 300 300
FT TURN 302 305
FT STRAND 313 313
FT STRAND 317 317
FT STRAND 321 323
FT HELIX 335 346
FT TURN 348 349
FT STRAND 358 370
FT TURN 371 372
FT STRAND 373 387
SQ SEQUENCE 390 AA; 44341 MW; 75391614250288FE CRC64;

```

Query Match 89.3%; Score 1936.5; DB 1; Length 390;  
Best Local Similarity 90.2%; Pred. No. 8.9e-148;  
Matches 367; Conservative 10; Mismatches 13; Indels 17; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI R G Q I L S K L R L A	60
Db	1	MPPSGLRLLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI R G Q I L S K L R L A	60
Qy	61	SPPSQGDVP PG PLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPG PLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVL MVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:   :   :	
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAHCSCD SKDNTLHVEINGFN	240
		:                     :         :	
Db	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWLSRGGEIEGFRLSAHCSCD SRDNTLQVDINGFT	240
Qy	241	SGRRGDLATI HG M NRPFL LL MATPLERA QHLHSSRHRRALDTNSYPYDV PDYASLA LD TN	300
		:	
Db	241	TGRRGDLATI HG M NRPFL LL MATPLERA OHLOSSRHRR-----ALDTN	283

```

QY      301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

QY      361 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      344 NQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 4

TGF1\_CERAE

```

ID      TGF1_CERAE      STANDARD;      PRT;      390 AA.
AC      P09533;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFB1.
OS      Cercopithecus aethiops (Green monkey) (Grivet).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC      Cercopithecinae; Cercopithecus.
OX      NCBI_TaxID=9534;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87246074; PubMed=3474130;
RA      Sharples K., Plowman G.D., Rose T.M., Twardzik D.R., Purchio A.F.;
RT      "Cloning and sequence analysis of simian transforming growth
RT      factor-beta cDNA.";
RL      DNA 6:239-244(1987).
CC      -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROL PROLIFERATION,
CC      DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
CC      CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM HAVE
CC      SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
CC      POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M16658; AAA35369.1; -.
DR      PIR; A26960; A26960.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.
DR      InterPro; IPR001839; TGFB.
DR      InterPro; IPR001111; TGFB_N.
DR      Pfam; PF00019; TGF-beta; 1.

```

DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbeta.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT PROPEP 17 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44356 MW; DFF63E2BAB44320E CRC64;

Query Match 89.2%; Score 1934.5; DB 1; Length 390;  
 Best Local Similarity 89.9%; Pred. No. 1.3e-147;  
 Matches 366; Conservative 10; Mismatches 14; Indels 17; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGRLRLPLLLPLLWLLVLT	TPSRPAAGLSTCKTIDMELVKRKRIE	TIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTR	DRVAGESAEPEPEPEADYYAKEVTR	VLMMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEP	VLLSRAELRLLRLKLKVEQHVELYQKYS	NDNSWR	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEP	VLLSRAELRLLRLKLKVEQHVELYQKYS	NNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWL	TREAI	EGFRLSAHCSCDSKDN	TLHVEINGFN 240
Db	181	YLSNRLLAPSNSPEWLSFDVTGVVRQWL	SRGGEI	EGFRLSAHCSCDSKDN	TLQVDINGFT 240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQ	HLHSSRHRRALDTNSYPYDVPDYASL	ALDTN	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQ	HLQSSRHRR-----	ALDTN	283
Qy	301	YCFSTEKNCCVRQLYIDFRKDLGWKWI	HEPKGYHANFCLGPCPYIWSLDTQYSK	VLALY	360
Db	284	YCFSTEKNCCVRQLYIDFRKDLGWKWI	HEPKGYHANFCLGPCPYIWSLDTQYSK	VLALY	343
Qy	361	NQHNPGASAAPCCVPQALEPLPIVYYV	GRKPKVEQLSNMIVRSCKCS	407	
Db	344	NQHNPGASAAPCCVPQALEPLPIVYYV	GRKPKVEQLSNMIVRSCKCS	390	

RESULT 5  
 TGF1\_CANFA  
 ID TGF1\_CANFA STANDARD; PRT; 390 AA.

AC P54831;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
GN TGFB1.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Jugular vein endothelial;  
RX MEDLINE=95237630; PubMed=7721110;  
RA Manning A.M., Auchampach J.A., Drong R.F., Slightom J.L.;  
RT "Cloning of a canine cDNA homologous to the human transforming growth  
RT factor-beta 1-encoding gene.";  
RL Gene 155:307-308(1995).  
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
CC -!- SUBUNIT: Homodimer; disulfide-linked.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the TGF-beta family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L34956; AAA51458.1; -.  
DR PIR; JC4023; JC4023.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR003911; TGF\_TGFB.  
DR InterPro; IPR001839; TGFB.  
DR InterPro; IPR001111; TGFB\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFB\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFB; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT PROPEP 24 278 BY SIMILARITY.  
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 285 294 BY SIMILARITY.  
FT DISULFID 293 356 BY SIMILARITY.  
FT DISULFID 322 387 BY SIMILARITY.

FT	DISULFID	326	389	BY SIMILARITY.
FT	DISULFID	355	355	INTERCHAIN.
FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL).
SO	SEQUENCE	390 AA; 44185 MW; EB4780E88B7B590E CRC64;		

Query Match 88.2%; Score 1913.5; DB 1; Length 390;  
Best Local Similarity 89.2%; Pred. No. 6.2e-146;  
Matches 363; Conservative 12; Mismatches 15; Indels 17; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
		:	
Db	1	MPPSGLRLLPLLLPLLRLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLS	60
Qy	61	SPPSQGDVPPGPPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPVPVPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVENTNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:     :     :	
Db	121	YEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRLKLKAEQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREA IEGFRLSAHCS CDSKDNTLHVEINGFN	240
		:                     : :                     :       :	
Db	181	YLSNRLLAPSDTPEWL SFDVTGVVRQWLSHGGEVEGFRLSAHCSCDSKDNTLQVDINGFS	240
Qy	241	SGRRGDLATI HGMNRPFL LLMATPLERA QHLHSSRRHRRALD TNSYPYDV PDYAS LALDTN	300
Db	241	SSRRGDLATI HGMNRPFL LLMATPLERA QHLHSSRQR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY	343
Qy	361	NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKEQLSNMIVRSCKCS	407
Db	344	NQHNPGASAAPCCVPOALEPLPIVYYVGRKPKEQLSNMIVRSCKCS	390

RESULT 6

```

TGF1_HORSE
ID      TGF1_HORSE      STANDARD;      PRT;      390 AA.
AC      O19011;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFBI.
OS      Equus caballus (Horse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX      NCBI_TaxID=9796;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Lymph node;

```





```

QY      61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 SPSPQGEVPPGPLPEAVLALYNSTRAQVAGESAETEPEPEADYYAKEVTRVLMVEKENEI 120

QY     121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVQHVELYQKYSNDSWR 180
      | : ||:| ||||| |||:|:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     121 YKTVETGSHSIYMFNTSELRAAVPDPMLLSRAELRLLRLKLSVEQHVELYQKYSNNSWR 180

QY     181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     181 YLSNRLLTPSDSPEWLSFDVTGVVRQWLSQGGAMEGFRLSAHCSCDSKDNTLRVGINGFS 240

QY     241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRHRRALDTNSYPYDVPDYASIALDTN 300
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     241 SSRRGDLATIDGMNRPFLLLMATPLERAQQLHSSRRHR-----ALDTN 283

QY     301 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     284 YCFSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

QY     361 NQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     344 NQHNPGASAAPCCVPQVLEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 7

## TGF1\_MOUSE

```

ID  TGF1_MOUSE      STANDARD;      PRT;      390 AA.
AC  P04202;
DT  20-MAR-1987 (Rel. 04, Created)
DT  20-MAR-1987 (Rel. 04, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Transforming growth factor beta 1 precursor (TGF-beta 1).
GN  TGFBI.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86168129; PubMed=3007454;
RA  Derynck R., Jarrett J.A., Chen E.Y., Goeddel D.V.;
RT  "The murine transforming growth factor-beta precursor.";
RL  J. Biol. Chem. 261:4377-4379(1986).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BALB/c;
RX  MEDLINE=96096545; PubMed=8522200;
RA  Guron C., Sudarshan C., Raghov R.;
RT  "Molecular organization of the gene encoding murine transforming
RT  growth factor beta 1.";
RL  Gene 165:325-326(1995).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6, and NOD/LT; TISSUE=Spleen;
RA  Poirrot L., Benoist C., Mathis D.;

```

```

RT  "Transforming growth factor-beta 1 sequence and expression: no
RT  difference between NOD/Lt and C57Bl/6 mouse strains.";
RL  Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M13177; AAA40423.1; -.
DR  EMBL; L42462; AAB00138.1; -.
DR  EMBL; L42456; AAB00138.1; JOINED.
DR  EMBL; L42457; AAB00138.1; JOINED.
DR  EMBL; L42458; AAB00138.1; JOINED.
DR  EMBL; L42459; AAB00138.1; JOINED.
DR  EMBL; L42460; AAB00138.1; JOINED.
DR  EMBL; L42461; AAB00138.1; JOINED.
DR  EMBL; AJ009862; CAA08900.1; -.
DR  PIR; A01396; WFMS2.
DR  HSSP; P01137; 1KLA.
DR  MGD; MGI:98725; Tgfb1.
DR  GO; GO:0005578; C:extracellular matrix; IDA.
DR  GO; GO:0006954; P:inflammatory response; IMP.
DR  GO; GO:0007515; P:lymph gland development; IMP.
DR  GO; GO:0008220; P:necrosis; IMP.
DR  GO; GO:0016202; P:regulation of myogenesis; IDA.
DR  InterPro; IPR002400; GF_cysknot.
DR  InterPro; IPR003911; TGF_TGfb.
DR  InterPro; IPR001839; TGfb.
DR  InterPro; IPR001111; TGfb_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGfb_propeptide; 1.
DR  PRINTS; PR00438; GFCYSKNOT.
DR  PRINTS; PR01423; TGFBETA.
DR  ProDom; PD000357; TGfb; 1.
DR  SMART; SM00204; TGFB; 1.
DR  PROSITE; PS00250; TGF_BETA_1; 1.
KW  Growth factor; Mitogen; Glycoprotein; Signal.
FT  SIGNAL          1      23      POTENTIAL.
FT  PROPEP          24      278
FT  CHAIN           279      390      TRANSFORMING GROWTH FACTOR BETA 1.
FT  DISULFID        285      294      BY SIMILARITY.
FT  DISULFID        293      356      BY SIMILARITY.
FT  DISULFID        322      387      BY SIMILARITY.
FT  DISULFID        326      389      BY SIMILARITY.

```

FT	DISULFID	355	355	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL).
SQ	SEQUENCE	390 AA;	44310 MW;	4381A51B7111D689E CRC64;

Query Match 85.5%; Score 1855.5; DB 1; Length 390;  
Best Local Similarity 85.5%; Pred. No. 2.8e-141;  
Matches 348; Conservative 15; Mismatches 27; Indels 17; Gaps 1;

Qy		1	MPSGLRLLPLLLLPLLWLVLTGPRPAAGLSTCKTIDMELVKKRIEAIRGQILSKRLA	60
Db		1	MPPSGLRLLPLLLLPLWLVLTGPRPAAGLSTCKTIDMELVKKRIEAIRGQILSKRLA	60
Qy		61	SPPSQGDVPFGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQI	120
			:                             :                     :	
Db		61	SPPSQGEVFPFPLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVL MVDNRNAI	120
Qy		121	YDKFKGTPHSLYMLFNNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
			:         :           :                             :	
Db		121	YEKTKDISHSIYMFFNNTSDIREAVPEPPLLSRAELRLQLRKSSVEQHVELYQKYSNNSWR	180
Qy		181	YLSNRLLAPSDSPDWLSFDVTGVVRQWL TRREAIEGFRLSAHCSCD SKDN TLHVEINGFN	240
			:   :                     : :   :                     :	
Db		181	YLG NRLL TPTDT PEWLSFDVTGVVRQWLNQG DGI Q GFR FSAHC SCDS KDN KL HVEINGIS	240
Qy		241	SGRRGDLATI HG MNRPFL LL MAT PLERAQHLHS SRHRRAL DT NSYPDY P DYAS L ALDTN	300
Db		241	PKRRGDLGTIHDMNRPFL LL MAT PLERAQHLHS SRHR-----ALDTN	283
Qy		301	YCFSSTEKNCCVRQLYIDFRKDLGWKIHEPKG YHANFCLGPCPYIW SLDTQYSKV LALY	360
Db		284	YCFSSTEKNCCVRQLYIDFRKDLGWKIHEPKG YHANFCLGPCPYIW SLDTQYSKV LALY	343
Qy		361	NQHNP GASAAPCCVPQA LEPLPI VYYVGRKPKEQ LSNMI VR SCKCS	407
			:	
Db		344	NQHNP GASASPCCVPQA LEPLPI VYYVGRKPKEQ LSNMI VR SCKCS	390

```

RESULT 8
TGF1_RAT
ID      TGF1_RAT          STANDARD;          PRT;      390 AA.
AC      P17246;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFB1.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Sprague-Dawley; TISSUE=Heart;
RX      MEDLINE=90272425; PubMed=2349108;

```



Qy	61	SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:                                     :                         :	
Db	61	SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:           :                                     :	
Db	121	YDKTKDITHSIYMFNTSDIREAVPEPPLLSRAELRLQRFKSTVEQHVELYQKYSNNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240
		:   :                         : :   :                     :	
Db	181	YLG NRLLTPTDTPEWLSFDVTGVVRQWLNQG DGI QGFRFSAHCS CDSKD NVLHVEINGIS	240
Qy	241	SGRRGDLATIHGMRPFLLL MATPLERAQHLHSSRHRRALDTNSYPYDVDPDYASLALDTN	300
Db	241	PKRRGDLGTIHDMNRP FLLL MATPLERAQHLHSSRHRR-----ALDTN	283
Qy	301	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD TQYSKVLALY	360
Db	284	YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD TQYSKVLALY	343
Qy	361	NQHNPGASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS	407
		:	
Db	344	NQHNPGASASPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS	390

TGF1 CAVPO

```

ID      TGF1_CAVPO          STANDARD;          PRT;      390 AA.
AC      Q9Z1Y6; Q9QZB3; Q9R148;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFB1.
OS      Cavia porcellus (Guinea pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX      NCBI_TaxID=10141;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Hartley;
RA      Jeevan A., McMurray D.N., Yoshimura T.;
RT      "Guinea pig transforming growth factor-beta in peritoneal exudates
RT      after BCG vaccination.";
RL      Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN      [2]
RP      SEQUENCE OF 265-382 FROM N.A.
RX      MEDLINE=99144670; PubMed=10025978;
RA      Scarozza A.M., Ramsingh A.I., Wicher V., Wicher K.;
RT      "Spontaneous cytokine gene expression in normal guinea pig blood and
RT      tissues.";
RL      Cytokine 10:851-859(1998).
RN      [3]
RP      SEQUENCE OF 279-371 FROM N.A.
RC      STRAIN=Hartley; TISSUE=Trachea;
RA      Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,

```

RA Sekizawa K.;  
 RT "Guinea-pig transforming growth factor-beta expression in injured  
 RT tracheal epithelium.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,  
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY  
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE  
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A  
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF191297; AAF02780.1; -.  
 DR EMBL; AF097509; AAC83807.1; -.  
 DR EMBL; AF169347; AAD49347.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT PROPEP 25 278 POTENTIAL.  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 FT CONFLICT 279 279 G -> P (IN REF. 3).  
 FT CONFLICT 286 286 F -> S (IN REF. 2).  
 FT CONFLICT 309 309 K -> E (IN REF. 2).  
 FT CONFLICT 322 322 C -> R (IN REF. 2).  
 FT CONFLICT 350 350 A -> G (IN REF. 2).  
 SQ SEQUENCE 390 AA; 44328 MW; 1539F849BA0C0FF1 CRC64;

Query Match 85.1%; Score 1846.5; DB 1; Length 390;  
 Best Local Similarity 85.3%; Pred. No. 1.5e-140;  
 Matches 347; Conservative 15; Mismatches 28; Indels 17; Gaps 1;

```

QY      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSRLRLLPLLLPLLWLLVLAPGRPASGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

QY     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEPEPDYYAKEVTRVLMVDNSHNI 120

QY    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YKSIETVAHSIYMFNTSELREAVDPDLLSRAELRMQRLKLNVEQHVELYQKYSNNSWR 180

QY    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNQLLTPSDTPEWLSFDVTGVVRQWLSQGEELEGFRFSAHCSCDSKDNTLRVEINGIG 240

QY    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLAATHGMNRPFLLLMATPLERAQHLHSSRHRR-----GLDTN 283

QY    301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    284 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343

QY    361 NQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    344 NQHNPASAAAPCCVPQALEPLPIVYYVGRKAKVEQLSNMIVRSCKCS 390
  
```

RESULT 10

TGF1\_BOVIN

```

ID   TGF1_BOVIN      STANDARD;          PRT;   315 AA.
AC   P18341;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1) (Fragment).
GN   TGFBI.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91042552; PubMed=3153459;
RA   van Obberghen-Schilling E., Kondaiah P., Ludwig R.L., Sporn M.B.,
RA   Baker C.C.;
RT   "Complementary deoxyribonucleic acid cloning of bovine transforming
RT   growth factor-beta 1.";
RL   Mol. Endocrinol. 1:693-698(1987).
RN   [2]
RP   SUBUNITS.
  
```



RC TISSUE=Bone;  
 RX MEDLINE=92129307; PubMed=1733936;  
 RA Ogawa Y., Schmidt D.K., Dasch J.R., Chang R.J., Glaser C.B.;  
 RT "Purification and characterization of transforming growth factor-beta  
 RT 2.3 and -beta 1.2 heterodimers from bovine bone.";  
 RL J. Biol. Chem. 267:2325-2328(1992).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked. Heterodimers of TGF-beta 1/2  
 CC have been found in bone.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M36271; AAA30778.1; -.  
 DR PIR; A40057; A40057.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT PROPEP <1 203  
 FT CHAIN 204 315 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 210 219 BY SIMILARITY.  
 FT DISULFID 218 281 BY SIMILARITY.  
 FT DISULFID 247 312 BY SIMILARITY.  
 FT DISULFID 251 314 BY SIMILARITY.  
 FT DISULFID 280 280 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 169 171 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 315 AA; 36269 MW; C2717A23D994E00E CRC64;

Query Match 73.4%; Score 1592.5; DB 1; Length 315;  
 Best Local Similarity 89.5%; Pred. No. 2.5e-120;  
 Matches 297; Conservative 9; Mismatches 9; Indels 17; Gaps 1;

Qy 76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135  
 | : | | | | | | | | | | | | | | | | | | | | | | | | | | | : | | | | | : | | : | | |  
 Db 1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMFF 60

Qy 136 NTSELREAVPEPVLLSRAELRLLRLKLVQHVLYQKYSNDSWRYLNSRLLAPSDSPEW 195  
 | | | | | | | | | | | | | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 61 NTSELREAVPEPVLLSRADVRLRLKLVQHVLYQKYSNNSWRYLNSRLLAPSDSPEW 120

Qy 196 LSFDTVGVVRQWLTRREAIEGFRLSAHCSCDSKDNLT LHVEINGFN SGRGDLATI HGMR 255  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 121 LSFDTVGVVRQWLTRREEIEGFRLSAHCSCDSKDNLT LQVDINGFSSGRGDLATI HGMR 180

Qy 256 PFLLLMATPLERAQHLHSSRHRALDTSYPYDVPDYASLALDTSYCFSSSTEKNCCVRQL 315  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 181 PFLLLMATPLERAQHLHSSRHR-----ALDTSYCFSSSTEKNCCVRQL 223

Qy 316 YIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVP 375  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 224 YIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVP 283

Qy 376 QALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 284 QALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 315

RESULT 11

TGF1\_CHICK

ID TGF1\_CHICK STANDARD; PRT; 373 AA.  
 AC P09531;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 4)  
 DE (Fragment).  
 GN TGFB1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White leghorn;  
 RX MEDLINE=89112198; PubMed=2464131;  
 RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;  
 RT "Complementary deoxyribonucleic acid cloning of a messenger  
 RT ribonucleic acid encoding transforming growth factor beta 4 from  
 RT chicken embryo chondrocytes.";  
 RL Mol. Endocrinol. 2:1186-1195(1988).  
 RN [2]  
 RP REVISIONS.  
 RX MEDLINE=92357039; PubMed=1353860;  
 RA Burt D.W., Jakowlew S.B.;  
 RT "Correction: a new interpretation of a chicken transforming growth  
 RT factor-beta 4 complementary DNA.";  
 RL Mol. Endocrinol. 6:989-992(1992).

```

CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M31160; AAB05637.1; -.
DR  PIR; A41918; A41918.
DR  HSSP; P01137; 1KLA.
DR  InterPro; IPR003911; TGF_TGFb.
DR  InterPro; IPR001839; TGFb.
DR  InterPro; IPR001111; TGFb_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGFb_propeptide; 1.
DR  PRINTS; PR01423; TGFbBETA.
DR  ProDom; PD000357; TGFb; 1.
DR  SMART; SM00204; TGFb; 1.
DR  PROSITE; PS00250; TGF_BETA_1; 1.
KW  Growth factor; Mitogen; Glycoprotein; Signal.
FT  NON_TER      1      1
FT  SIGNAL      <1      1      POTENTIAL.
FT  PROPEP       2     259      POTENTIAL.
FT  CHAIN       260    373      TRANSFORMING GROWTH FACTOR BETA 1.
FT  DISULFID    266    277      BY SIMILARITY.
FT  DISULFID    276    339      BY SIMILARITY.
FT  DISULFID    305    370      BY SIMILARITY.
FT  DISULFID    309    372      BY SIMILARITY.
FT  DISULFID    338    338      INTERCHAIN (BY SIMILARITY).
FT  CARBOHYD     54     54      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD    109    109      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD    153    153      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  SITE        224    226      CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE    373 AA;  42634 MW;  9903F3479C8552E5 CRC64;

```

```

Query Match          49.2%;  Score 1067;  DB 1;  Length 373;
Best Local Similarity 53.5%;  Pred. No. 4e-78;
Matches 209;  Conservative 52;  Mismatches 98;  Indels 32;  Gaps 8;

```

```

Qy      30 LSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVA 89
      |||:| :|:| | :|:||||:|:|:|:|:|:|:|:| :| :| |||: | |||:|:|:|
Db      2 LSTCQRLDLEAAKKKRIEAVRGQILSKLRLTAPPPASETPPRPLPDDVRALYNSTQELLK 61

Qy      90 GES-VEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148
      : : | |:| :|:|:|:| :| :| : : : :|:|:|:| | |
Db      62 QRARLRPPPDGPDEYWAKELRRIPMETTWGAMEHWQPQSHSIFVFNVSRARRG-GRPT 120

```

```

QY      149 LLSRAELRLLRLKLK-----VEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGV 203
      || |||||:| | || :||| | | |||| | : : |||||
Db      121 LLHRAELRMLRQKAAADSAGTEQRLLEYQGYGNASWRYLHGRSVRATADDEWLSFDVTD 180

QY      204 VRQWLTRREAIEGFRLSAHCSCD---SKDNTLHVEINGFNSGRRGDLATIHGMNR--PFL 258
      | |||: | : |:|| | | : : | | :|||: | : | |::
Db      181 VHQWLSGSELLGVFKLSVHCPCMGPGHAEEMRISIEGFQ-QRGDMQSIKKHRRVPYV 239

QY      259 LLMATPLERAQHLHSSRRRRALDTSNYPYDVPDYASLALDTNYCF--SSTEKNCCVRQLY 316
      | || | || | |||:| || |||:||| : ||||| ||
Db      240 LAMALPAERANELHSARRRR-----DLDTDYCFGPGTDEKNCCVRPLY 282

QY      317 IDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQ 376
      ||||| ||||| |||||:||||| ||||:||||| ||||| |||||
Db      283 IDFRKDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNPGASAAPCCVPQ 342

QY      377 ALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
      |:||||:||||| :|||||:|:||||
Db      343 TLDPLPIIYVGRNVRVEQLSNMVVRACKCS 373

```

# RESULT 12

TGF1\_XENLA

ID TGF1\_XENLA STANDARD; PRT; 382 AA.

AC P16176;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 5).

OS *Xenopus laevis* (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

OC Xenopodinae; *Xenopus*.

OX NCBI\_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90110090; PubMed=2295601;

RA Kondaiah P., Sands M.J., Smith J.M., Fields A., Roberts A.B.,

RA Sporn M.B., Melton D.A.;

RT "Identification of a novel transforming growth factor-beta (TGF-beta

5) mRNA in *Xenopus laevis*.";

RL J. Biol. Chem. 265:1089-1093(1990).

RN [2]

RP SEQUENCE FROM N.A.

RA Vempati U.D., Kondaiah P.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: IMPORTANT ROLE IN CERTAIN ASPECTS OF DIFFERENTIATION.

CC -!- SUBUNIT: Homodimer; disulfide-linked.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the TGF-beta family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; J05180; AAA49968.1; -.  
DR EMBL; AF009335; AAB64441.1; -.  
DR EMBL; AF009331; AAB64441.1; JOINED.  
DR EMBL; AF009332; AAB64441.1; JOINED.  
DR EMBL; AF009333; AAB64441.1; JOINED.  
DR EMBL; AF009334; AAB64441.1; JOINED.  
DR PIR; A34929; B61036.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR003911; TGF\_TGFB.  
DR InterPro; IPR001839; TGFB.  
DR InterPro; IPR001111; TGFB\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFB\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFB; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT PROPEP 22 270  
FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 277 286 BY SIMILARITY.  
FT DISULFID 285 348 BY SIMILARITY.  
FT DISULFID 314 379 BY SIMILARITY.  
FT DISULFID 318 381 BY SIMILARITY.  
FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 234 236 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 382 AA; 44200 MW; 1034621C917AAE15 CRC64;

Query Match 40.9%; Score 887.5; DB 1; Length 382;  
Best Local Similarity 46.4%; Pred. No. 1e-63;  
Matches 192; Conservative 54; Mismatches 121; Indels 47; Gaps 11;

Qy 9 LPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDV 68  
: :| || |||| | |||| :||| |::| ||||| ||||| :| |  
Db 1 MEVLWMLLVLLVLHLSSLAMSLSTCKAVDMEEVRKRRIEAI RQILSKLKLDKTPDV-DS 59  
  
Qy 69 PPGPLPEAVLALYNSTRDRVAGESVEPE-----PEPEADYYAKEVTRVLMVESGNQIYDK 123  
:| : ||||| : : : :| : ||||| :| | :| :|  
Db 60 EKMTVPSEAI FLYNSTLEVIREKATREEEHVGH DQNIQDY YAKQVYR---FESITELED- 115  
  
Qy 124 FKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKL--KVEQHVELYQKYSNDSW-- 179  
| ||| :||| || ||||| : : : :| :||| :| :  
Db 116 -----HEFKFKFNASHVRENVGMNSLLHHAEL RMYKKQTDKNMDQRMELFWKYQENGTH 170  
  
Qy 180 -RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC---DSKDNTLHVE 235  
||| : : | ||:|||| | :|| | | | | | :|| : :  
Db 171 SRYLESKYITPVTDDWMSFDVTKT VNEWLKRAEENEQFGLQPACKCPTPQAKD----ID 226

Qy	236	INGFNSGRGDLATI HGM--NRPFLLLMATPLERAQHLHSSRRHRRALD	293
		: : :     :     :	
Db	227	IEGF-PALRGDLASLSSKENTKPYLMITSMIPAERIDTVTSSRKKR-----	270
Qy	294	SLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQY	353
		:     :     :     :                       :	
Db	271	--GVGQEYCFGNNGPNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQY	328
Qy	354	SKVLALYNQHNP GASAAPCCVPQALEPLPIVYVVG RKPKEQLSNMIVRSCKCS	407
		:       :     :     :	
Db	329	SKVLSLYNONNPGASISPCVDPDVL EPLPIIYVVGRTAKVEQLSNMIVRSCNCS	382

## TGF3 CHICK

```

AC      P16047;
DT      01-APR-1990 (Rel. 14, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 3 precursor (TGF-beta 3).
GN      TGFB3.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89096966; PubMed=3211158;
RA      Jakowlew S.B., Dillard P.J., Kondaiah P., Sporn M.B., Roberts A.B.;
RT      "Complementary deoxyribonucleic acid cloning of a novel transforming
RT      growth factor-beta messenger ribonucleic acid from chick embryo
RT      chondrocytes.";
RL      Mol. Endocrinol. 2:747-755(1988).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=White leghorn;
RX      MEDLINE=95169270; PubMed=7865129;
RA      Burt D.W., Dey B.R., Paton I.R., Morrice D.R., Law A.S.;
RT      "The chicken transforming growth factor-beta 3 gene: genomic
RT      structure, transcriptional analysis, and chromosomal location.";
RL      DNA Cell Biol. 14:111-123(1995).
RN      [3]
RP      SEQUENCE OF 1-117 FROM N.A.
RC      STRAIN=White leghorn; TISSUE=Blood;
RX      MEDLINE=92134496; PubMed=1840616;
RA      Burt D.W., Dey B.R., Paton I.R.;
RT      "Comparative analysis of human and chicken transforming growth
RT      factor-beta 2 and -beta 3 promoters.";
RL      J. Mol. Endocrinol. 7:175-183(1991).
RN      [4]
RP      SEQUENCE OF 1-117 FROM N.A.
RX      MEDLINE=93024487; PubMed=1406706;
RA      Jakowlew S.B., Lechleider R., Geiser A.G., Kim S.J.,
RA      Santa-Coloma T.A., Cubert J., Sporn M.B., Roberts A.B.;
RT      "Identification and characterization of the chicken transforming

```



Qy 75 EAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVE---SGNQIYDKF 124  
 :|||||: : | :| | | | :|||: : | :| :| :  
 Db 67 YQILALYNSTRELL--EEMEEEEKEESCSQENTESEYYAKEIHKFDMIQGLPEHNELGICP 124

Qy 125 KGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND--- 177  
 || : : | | | | | | | :||| : || :||| : |  
 Db 125 KGVTSNVFR-FNVS---SAEKNSTNLFRAEFRVLRVPNPSSKRSEQRIELFQILRPDEHI 180

Qy 178 -SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC-----D 226  
 |||| : | ||||| ||| : || : | | : :  
 Db 181 AKQRYLSGRNVQTRGSPEWLSFDVTDTVREWLLHRESNLGLEISIHCPCHTFQPNGDILE 240

Qy 227 SKDNTLHVEINGFNSG---RRGDLATI---HGMNRPFLLLMATPLERAQH--LHSSRHRR 278  
 : | : : | : | | : : | :| | | : | :| :|  
 Db 241 NLHEVLEIKFKGIDSEDDYGRGDLGRLKKQKDLHNPHLILMMLPPHRLESPTLGGQRKKR 300

Qy 279 ALDTNSYPYDVPDYASLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANF 338  
 ||||| : | :||| | | : ||||| : ||||| |||  
 Db 301 -----ALDTNYCFRNLEENCCVRPLYIELPTDLGWKWVHEPKGYFANF 343

Qy 339 CLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSN 398  
 | |||| : | || : | | || | | ||| : |||| | ||||| |||||  
 Db 344 CSGPCPYLRSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSN 403

Qy 399 MIVRSCKCS 407  
 | : | : |||||  
 Db 404 MVVKSCCKCS 412

# RESULT 14

## TGF3\_PIG

ID TGF3\_PIG STANDARD; PRT; 409 AA.  
 AC P15203;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 3 precursor (TGF-beta 3).  
 GN TGFB3.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=89091120; PubMed=3208746;  
 RA Derynck R., Lindquist B., Lee A., Wen D., Tamm J., Graycar J.L.,  
 RA Rhee L., Mason A.J., Miller D.A., Coffey R.J., Moses H.L.,  
 RA Chen E.Y.;  
 RT "A new type of transforming growth factor-beta, TGF-beta 3."  
 RL EMBO J. 7:3737-3743(1988).  
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X14150; CAA32363.1; -.  
DR PIR; S01825; S01825.  
DR HSSP; P10600; 1TGJ.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR003911; TGF\_TGFB.  
DR InterPro; IPR001839; TGFB.  
DR InterPro; IPR001111; TGFB\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFB\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFB; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 297  
FT CHAIN 298 409 TRANSFORMING GROWTH FACTOR BETA 3.  
FT DISULFID 304 313 BY SIMILARITY.  
FT DISULFID 312 375 BY SIMILARITY.  
FT DISULFID 341 406 BY SIMILARITY.  
FT DISULFID 345 408 BY SIMILARITY.  
FT DISULFID 374 374 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 259 261 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 409 AA; 46814 MW; B4900235B5CC955E CRC64;

Query Match 39.2%; Score 851; DB 1; Length 409;  
Best Local Similarity 44.6%; Pred. No. 9.3e-61;  
Matches 190; Conservative 58; Mismatches 122; Indels 56; Gaps 13;

Qy 15 LLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGGLP 74  
|:|:| : :||| |:| : :|||:|||||:||||| : :|  
Db 7 LVLALLNFATVSLSMSTCTTLDHDKRKRVIAHQILSKLRLTSPDPMSML--ANIP 64  
  
Qy 75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMV---ESGNQIYDKFKG 126  
|| ||||| : | || : : :|||:| : | : | : ||  
Db 65 TQVLDLYNSTRELLEEVHGERGDDCTQENTESEYYAKEIYKFDMIQGLEEHNDLAVCPKG 124  
  
Qy 127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND---S 178  
: : || | : : | ||| |:|: : || :||:| |  
Db 125 ITSKIFR-FNVSSVEK---NETNLFRAEFRLRMPNPSSKRSEQRIELFQILQPDEHIAK 180  
  
Qy 179 WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC-----DSK 228  
||: : | : ||||| |:| |||: | :| ||| : :  
Db 181 QRYIDGKNLPTRGAAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENI 240

```

Qy      229 DNTLHVEINGFNS---GRRGDLATIHGM--NRPFLLLMATPLERAQH--LHSSRHRRALD 281
      :  ::  |  :|      ||||  :      :  |  :||  |  :|  :  |  :  |
Db      241 QEVM EIKFKGVDS EDDPGRGDLGRLKKKKEHSPHLILMMIPDRLDNPG LG AQRKKR--- 297

Qy      282 TNSYPYDVPDYASLALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLG 341
      |||||  |||||  :  |  :|||  |||||  :|||  :|||  :|||  :|||  |
Db      298 -----ALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQWVHEPKGYANFCSG 343

Qy      342 PCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIV 401
      ||||  :  ||  :|  ||  ||  ||  :|||  |||  |  :|||  |||||  :|
Db      344 PCPYLRSADTTHSSVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTAKVEQLSNMVV 403

Qy      402 RSCKCS 407
      :|||
Db      404 KSCKCS 409

```

# RESULT 15

## TGF3\_MOUSE

```

ID      TGF3_MOUSE      STANDARD;      PRT;      410 AA.
AC      P17125;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 3 precursor (TGF-beta 3).
GN      TGFB3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90190650; PubMed=2628730;
RA      Miller D.A., Lee A., Matsui Y., Chen E.Y., Moses H.L., Derynck R.;
RT      "Complementary DNA cloning of the murine transforming growth
RT      factor-beta 3 (TGF beta 3) precursor and the comparative expression
RT      of TGF beta 3 and TGF beta 1 messenger RNA in murine embryos and
RT      adult tissues.";
RL      Mol. Endocrinol. 3:1926-1934(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91000714; PubMed=2206556;
RA      Denhez F., Lafyatis R., Kondaiah P., Roberts A.B., Sporn M.B.;
RT      "Cloning by polymerase chain reaction of a new mouse TGF-beta,
RT      mTGF-beta 3.";
RL      Growth Factors 3:139-146(1990).
CC      -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.

```

```

CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M32745; AAA40422.1; -.  
 DR PIR; A41397; A41397.  
 DR HSSP; P10600; 1TGJ.  
 DR MGD; MGI:98727; Tgfb3.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.

FT	SIGNAL	1	23	POTENTIAL.
FT	PROPEP	24	298	POTENTIAL.
FT	CHAIN	299	410	TRANSFORMING GROWTH FACTOR BETA 3.
FT	DISULFID	305	314	BY SIMILARITY.
FT	DISULFID	313	376	BY SIMILARITY.
FT	DISULFID	342	407	BY SIMILARITY.
FT	DISULFID	346	409	BY SIMILARITY.
FT	DISULFID	375	375	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD	72	72	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	133	133	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	259	261	CELL ATTACHMENT SITE (POTENTIAL).
SQ	SEQUENCE	410 AA; 46884 MW; 250F7048CA432BD6 CRC64;		

Query Match 38.9%; Score 844.5; DB 1; Length 410;  
 Best Local Similarity 44.2%; Pred. No. 3.1e-60;  
 Matches 188; Conservative 57; Mismatches 127; Indels 53; Gaps 12;

Qy	15	LLWLLVLT	PGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLASPPSQGDVPPG	PLP	74
		: :	: : : : : : : : : : :	: : : : : : : : : : :	:	
Db	7	LVVLALLNL	ATISLSLSTCTTLD	FGHIKKRVEAIRGQILSKLRLTSP	PEPSVMT--HVP	64
Qy	75	EAVLALYNSTR--	DRVAGESVE--PEPEPEADYYAKEVTR	VLMEV--SGNQIYDKFKG		126
			: : : : : : : : : : :	: : : : : : : : : : :	:	
Db	65	YQVLALYNSTREL	LEEMHGEREEGCTQETSESEYYAKEIHK	FDMIQGLAEHNELAVCPKG		124
Qy	127	TPHSLYMLFNTSEL	REAVPEPVLLSRAELRLLRL----	KLKVEQHVELYQKYSND----	S	178
		: : :	: : : : : : : : : : :	: : : : : : : : : : :	:	
Db	125	ITSKVFR-FNVSS	VEK--NGTNLFRAEFRVLRVPNPSSKR	TEQRIELFQILRPDEHIAK		180
Qy	179	WRYLSNRLLAPSDS	PEWLSFDVTGVVRQWLTRREAIEGFR	LSAHCSC-----DSK		228
		: :	: : : : : : : : : : :	: : : : : : : : : : :	:	
Db	181	QRYIGGKNLPTRG	TAEWLSFDVTDTVREWLLRRESNLGLE	ISIHCPCHTFQPNGDILENV		240
Qy	229	DNTLHVEINGFNS--	GRRGDLATIHGM--NRPFLLMATPLER	AQHLHSSRHRRALDT		282
		: : :	: : : : : : : : : : :	: : : : : : : : : : :	:	
Db	241	HEVMEIKFKGVDN	EDDHGRGDLGRLKKQKDHHPHILMMI	PPHRLDSPGQGSQRK----		296
Qy	283	NSYPYDVPDYASL	ALDTNYCFSSTEKNCCVRQLYIDFRK	DLGWKWIHEPKGYHANFCLGP		342

```

          ||||| : ||||| |||||:|||||:|||||:|||||
Db      297 -----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKVVHEPKGYANFCSGP 345
Qy      343 CPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVR 402
          |||: |||: ||| ||| ||| |||:||||| ||| |:||||| ||||| |||:|:
Db      346 CPYLRSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVK 405
Qy      403 SCKCS 407
          |||||
Db      406 SCKCS 410

```

Search completed: October 28, 2003, 09:08:42  
Job time : 9.15764 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 07:50:55 ; Search time 32.6306 Seconds  
(without alignments)  
3218.683 Million cell updates/sec

Title: US-10-017-372E-39  
Perfect score: 2169  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 407

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*

```

14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	1835.5	84.6	390	6	Q9TUM8	Q9tum8 equus cabal	
2	1756.5	81.0	368	11	Q8R4D9	Q8r4d9 sigmodon hi	
3	868.5	40.0	379	13	Q8JHF5	Q8jhf5 sparus aura	
4	864.5	39.9	379	13	Q8AXK8	Q8axk8 sparus aura	
5	844.5	38.9	412	11	Q91YU7	Q91yu7 mus musculu	
6	843	38.9	382	13	Q9PWA9	Q9pwa9 morone chry	
7	841	38.8	414	11	Q91VP5	Q91vp5 mus musculu	
8	826	38.1	382	13	O93449	O93449 oncorhynchu	
9	787	36.3	376	13	Q9PTQ2	Q9ptq2 cyprinus ca	
10	780	36.0	399	11	Q9ERB7	Q9erb7 mesocricetu	
11	736.5	34.0	362	11	Q99K17	Q99k17 mus musculu	
12	693.5	32.0	361	13	Q98854	Q98854 cyprinus ca	
13	691.5	31.9	130	11	Q08714	Q08714 mesocricetu	
14	674.5	31.1	124	6	Q95N80	Q95n80 canis famil	
15	638	29.4	112	6	O02730	O02730 oryctolagus	
16	620	28.6	255	11	Q921T1	Q921t1 mus musculu	
17	604.5	27.9	127	6	Q9TV08	Q9tv08 canis famil	
18	590	27.2	224	11	Q8CDZ9	Q8cdz9 mus musculu	
19	580.5	26.8	200	13	Q90YF1	Q90yf1 pleuronecte	
20	554	25.5	101	11	Q9R184	Q9r184 meriones un	
21	476	21.9	179	13	Q90YF2	Q90yf2 pleuronecte	
22	413	19.0	88	13	Q90YF5	Q90yf5 pleuronecte	
23	402	18.5	88	13	Q90YF7	Q90yf7 oncorhynchu	
24	397	18.3	88	13	Q90ZE7	Q90ze7 acipenser b	
25	393	18.1	87	13	O42306	O42306 carassius a	
26	388.5	17.9	309	4	Q8WV88	Q8wv88 homo sapien	
27	383	17.7	91	6	Q9MYZ1	Q9myz1 capra hircu	
28	373	17.2	86	6	Q28241	Q28241 cervus elap	
29	358	16.5	81	6	Q9N1S3	Q9nls3 capreolus c	
30	340	15.7	87	13	Q8JHB6	Q8jhb6 scophthalmu	
31	336.5	15.5	375	13	Q8UWD8	Q8uwd8 columba liv	
32	323.5	14.9	375	13	Q8AVB2	Q8avb2 coturnix co	
33	321.5	14.8	375	13	Q8UWD7	Q8uwd7 coturnix ch	
34	321.5	14.8	389	13	Q90YY0	Q90yy0 ictalurus p	
35	317	14.6	77	13	Q90YF8	Q90yf8 oncorhynchu	
36	316.5	14.6	375	13	Q98SP0	Q98sp0 gallus gall	
37	313.5	14.5	375	6	Q9GM97	Q9gm97 equus cabal	
38	313.5	14.5	375	13	Q8UWE0	Q8uwe0 anas platyr	
39	311.5	14.4	375	13	Q8UWD9	Q8uwd9 anser anser	
40	310.5	14.3	375	6	Q8HY52	Q8hy52 lepus capen	
41	307.5	14.2	375	6	Q95J86	Q95j86 macaca fasc	
42	302	13.9	50	6	Q28240	Q28240 cervus elap	
43	301	13.9	62	13	Q90ZJ7	Q90zj7 anguilla an	

44	300	13.8	62	13	Q90YF4	Q90yf4 pleuronecte
45	296.5	13.7	375	6	Q8WNS6	Q8wns6 bos taurus

# ALIGNMENTS

## RESULT 1

### Q9TUM8

ID Q9TUM8 PRELIMINARY; PRT; 390 AA.  
AC Q9TUM8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Transforming growth factor beta 1.  
GN TGFBI.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Nixon A.J., Brower-Toland B.T., Sandell L.J.;  
RT "Molecular cloning of equine transforming growth factor beta 1 reveals  
RT a divergent nucleotide structure that encodes a novel bioactive  
RT peptide among mammalian species."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR EMBL; AF175709; AAD49431.1; -.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFbeta.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
SQ SEQUENCE 390 AA; 43860 MW; 220FE40DFCCA6016 CRC64;

Query Match 84.6%; Score 1835.5; DB 6; Length 390;  
Best Local Similarity 86.0%; Pred. No. 2e-155;  
Matches 350; Conservative 12; Mismatches 28; Indels 17; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTR	AQVAGESAETEPEPEADYYAKEVTR	VLMMVEKENEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQHVELY	QKYSNDSWR	180

```

Db      121 YKTVETGSHSIYMFFNASELRRAAVPDPMLLSRAELRLLRLKLSVEQHVELYQKYSNNSWR 180
Qy      181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 YLSNRLLTPSDSPEWLSFDVTGVVRQWLSQGGAMEGLRLSAHCPDCSKDNTLRVGINGFS 240
Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALDTN 300
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 SSRRGDLATIDGMNRPFLLLMATPLERAQQLHSSRHRR-----ALDTN 283
Qy      301 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 360
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      284 YCSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 343
Qy      361 NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      344 NQHNPASAAAPCCVPQVLEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 2

Q8R4D9

```

ID      Q8R4D9      PRELIMINARY;      PRT;      368 AA.
AC      Q8R4D9;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta-1 protein (Fragment).
GN      TGFb1.
OS      Sigmodon hispidus (Hispid cotton rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC      Sigmodon.
OX      NCBI_TaxID=42415;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Blanco J.C., Pletneva L.M., Prince G.A.;
RT      "Cotton rat cytokines, chemokines, and interferons.";
RL      Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AF480858; AAL87199.1; -.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFb; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
FT      NON_TER      1      1
SQ      SEQUENCE      368 AA; 41905 MW; A5C91207B0468B4A CRC64;

```

```

Query Match      81.0%; Score 1756.5; DB 11; Length 368;
Best Local Similarity 84.9%; Pred. No. 2.1e-148;
Matches 327; Conservative 16; Mismatches 25; Indels 17; Gaps 1;

```

Qy 23 PGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVPPGPLPEAVLALYN 82

```

      |||
Db      1 PGRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKLRLASPPSQGEVPPGPLPEAVLALYN 60
Qy      83 STRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELRE 142
      |||
Db      61 STRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAIYDKTKDIPHSVYMFNTSDIRE 120
Qy      143 AVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTG 202
      |||
Db      121 AVPEPPLLSRAELRLQRFKSNVEQHVELYEKYSNNSWRYLGNRLLSPTDSPEWLSFDVTS 180
Qy      203 VVRQWLTRREAIEGFRLSAHCSCDSKDNTHLVEINGFNSGRRGDLATIHGMNRPFLLLMA 262
      |||
Db      181 VVRKWLNQGDGIQGFRFSAHCSCDSKDNILHVEINGISPKRRGDLGTIHDNMRPFLLLMA 240
Qy      263 TPLERAQHLHSSRHRRALDTSYPYDVPDYASLALDTSNYCFSSTEKNCCVRQLYIDFRKD 322
      |||
Db      241 TPLERAQHLHSSRHR-----ALDTSNYCFSSTEKNCCVRQLYIDFRKD 283
Qy      323 LGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLP 382
      |||
Db      284 LGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASASPCCVPALEPLP 343
Qy      383 IVYYVGRKPKVEQLSNMIVRSCKCS 407
      |||
Db      344 IVYYVGRKPKVEQLSNMIVRSCKCS 368

```

# RESULT 3

Q8JHF5

```

ID      Q8JHF5          PRELIMINARY;          PRT;    379 AA.
AC      Q8JHF5;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta 1.
OS      Sparus aurata (Gilthead sea bream).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC      Sparidae; Sparus.
OX      NCBI_TaxID=8175;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Tafalla C., Aranguren R., Secombes C.J., Castrillo J.L., Novoa B.,
RA      Figueras A.;
RT      "Molecular characterization of sea bream (Sparus aurata) transforming
RT      growth factor beta1.";
RL      Fish and Shellfish Immunol. 0:0-0(2002).
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AF510084; AAN03842.1; -.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      InterPro; IPR003911; TGF_TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 2.

```



Query Match 40.0%; Score 868.5; DB 13; Length 379;  
Best Local Similarity 46.2%; Pred. No. 4.5e-69;  
Matches 189; Conservative 71; Mismatches 104; Indels 45; Gaps 16;

```

RESULT 4
Q8AXK8
ID   Q8AXK8          PRELIMINARY;          PRT;    379 AA.
AC   Q8AXK8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1.
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tafalla C., Novoa B., Aranuren R., Fiqueras A.;

```

RT "Molecular cloning and characterization of sea bream (*Sparus aurata*)  
RT TGF beta 1.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF424703; AAN76665.1; -.  
SQ SEQUENCE 379 AA; 43485 MW; A306EC387F6DBA7C CRC64;

Query Match 39.9%; Score 864.5; DB 13; Length 379;  
Best Local Similarity 46.2%; Pred. No. 1e-68;  
Matches 189; Conservative 70; Mismatches 105; Indels 45; Gaps 16;

```

Qy      12 LLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRL-ASPPSQGDVPP 70
      |:|:::| |: :|:||||:|:|:||||| | | | | | | | | | | | | | | | | |
Db      3 LVFLMFMVAYTVGK-VSGMSTCKTLDLEMVKKKRIEAIHQILSKLRLPTESPQAGD--E 59

Qy      71 GPLPEAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVESGNQIYDKFK 125
      :| :|:|:| | | | : : : | : : | :|:| | | | : | :
Db     60 EEIPSSLLSLYNSTKEMLKEQQTEVQTDIFTEXEEEEEYFA---TRVHKFNNTNPV----- 111

Qy     126 GTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKV--EQHVELYQKYSNDSWRYL 182
      || :| | | | | : : : ||: |||:| :| | | | | | | | | | | | | |
Db    112 RTPQNMSMSFNISEIRRSIGDYRLTLTAELRML-IKAPTILDEQRVELYQGLGT-SPRYL 169

Qy     183 SNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAHCSC--DSKDNTLHVEINGFN 240
      ::| : :| | | | | : : || : : : | : | :| | | :| :|
Db    170 ASRFITNELRDKWLSFDVTETLQNLKGNDDVQVFQLRLYCDCGRSSDVSTFSFGISGMT 229

Qy     241 SGRRGDLATIHGMNR--PFLLLMATPLERAQHLHSSRHRRALDTNSYPYDVPDYASLALD 298
      :| ||| | : | : | : | : || :|| :| :| :| :|
Db    230 AG-RGDKAVLDDMTKQPPYILTMSIPKNVSSHL-TSRKKRSTETK-----D 273

Qy     299 TNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLA 358
      | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    274 T--CTAQTE-TCCVRSLYIDFRKDLGWKWIHKPTRYHANYCMGSCTYIWNNAENKYSQXLA 330

Qy     359 LYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    331 LYKHHNPGASAPCCVPQALEPLPIVYVGRQHKVEQLSNMIVKSKCS 379

```

# RESULT 5

Q91YU7

ID Q91YU7 PRELIMINARY; PRT; 412 AA.

AC Q91YU7;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Transforming growth factor, beta 3.

GN TGFB3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.

CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; BC014690; AAH14690.1; -.  
DR MGD; MGI:98727; Tgfb3.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFbBETA.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
SQ SEQUENCE 412 AA; 47144 MW; F3EB65D046DF32AD CRC64;

### RESULT 6

ID Q9PWA9 PRELIMINARY; PRT: 382 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta precursor.  
 GN TGF-BETA.  
 OS Morone chrysops x Morone saxatilis (white bass x striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;  
 OC Moronidae; Morone.  
 OX NCBI\_TaxID=45352;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=KIDNEY;  
 RX MEDLINE=20394636; PubMed=10938723;  
 RA Harms C.A., Kennedy-Stoskopf S., Horne W.A., Fuller F.J.,  
 RA Tompkins W.A.F.;  
 RT "Cloning and sequencing hybrid striped bass (Morone saxatilis x M.  
 RT chrysops) transforming growth factor-beta (TGF-beta), and development  
 RT of a reverse transcription quantitative competitive polymerase chain  
 RT reaction (RT-qcPCR) assay to measure TGF-beta mRNA of teleost fish.";  
 RL Fish Shellfish Immunol. 10:61-85(2000).  
 CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE  
 CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- TISSUE SPECIFICITY: HIGHER LEVELS FOUND IN MONONUCLEAR CELLS FROM  
 CC PERIPHERAL BLOOD THAN IN SPLEEN OR ANTERIOR KIDNEY.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF140363; AAD46997.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 ? POTENTIAL.  
 FT PROPEP ? 270  
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.  
 FT DISULFID 278 286 BY SIMILARITY.  
 FT DISULFID 285 348 BY SIMILARITY.  
 FT DISULFID 314 379 BY SIMILARITY.  
 FT DISULFID 318 381 BY SIMILARITY.  
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 235 237 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 382 AA; 43846 MW; 3124D8C34EA74D72 CRC64;

Query Match 38.9%; Score 843; DB 13; Length 382;  
 Best Local Similarity 45.0%; Pred. No. 8.5e-67;  
 Matches 185; Conservative 67; Mismatches 107; Indels 52; Gaps 14;

```

Qy      15 LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRL--ASPPSQGDVPPG 71
      | : : : | | : : | | | | : : | : | | | | | | | | | | | | | |
Db      6 LMLVVVYTVGN-VSGMSTCKTLDLEMVKKKRIEAI RSQILSKLRLPKPEPDQAGDEEEI 64

Qy     127 PLPEAVLALYNSTRDRVAGESVEPE-----PEPEADYYAKEVTRVLMVESGNQIYDKFKG 126
      | | : : | | | | : : : | : : : | : : | | : : | |
Db     65 PTP--LLSLYNSTKEMLKEQQTEVQTDISTEQEEEEYFAKVLHKFNMTRKNN----- 114

Qy     127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKV--EQHVELYQKYSNDSWRYLSN 184
      | : | | | | : : | : | : | | | : | | : | | | | | | | | | : :
Db    115 TDTTKKMFFNISEIRESVGDYRLLTSAELRMLIKKTTIYDEQRVELYSGL-GDSPRYLAS 173

Qy     185 RLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHCSCD---SKDNTLHVEINGFNS 241
      | : : : | | | | : : | | : : | : | | | | : : : : | : | :
Db    174 RFITNKWKDKWLSFDVTKTLQDWLKGTDDEQGFQLRLFCCECNKVSAGETIFKFGISGIDP 233

Qy     242 GRRGDLATIHGMNR--PFLLLMATPLERAQHLHS---SRHRRALDTSYPYDVPDYASLA 296
      | | | | : : : | : | | : | | : | | : | : | : |
Db    234 G-RGDTGPMQLLTQQPPYILTMSIP---QNISSPSTSRKKRSTETK----- 275

Qy     297 LDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKV 356
      : | : | | | | | | | | | | | | | | | | | | | | | | | : : | : :
Db    276 ---DVCTAQTE-TCCVRSLYIDFRKDLGWKWIHKPTGYNANYCMGSCTYIWNNAENKYSQI 331

Qy     357 LALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 407
      | | | | | | | | | | | | | | | : | | | : | | | | | | | | | |
Db    332 LALYKHHNPGASAPCCVPQALEPLPILYVGRQHKEQLSNMIVKSKCS 382

```

# RESULT 7

Q91VP5

```

ID   Q91VP5          PRELIMINARY;          PRT;      414 AA.
AC   Q91VP5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to transforming growth factor, beta 2.
GN   TGFB2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; BC011170; AAH11170.1; -.
DR   MGD; MGI:98726; Tgfb2.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.

```

DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 SQ SEQUENCE 414 AA; 47588 MW; DB37A7C38881F286 CRC64;

Query Match 38.8%; Score 841; DB 11; Length 414;  
 Best Local Similarity 43.8%; Pred. No. 1.4e-66;  
 Matches 190; Conservative 62; Mismatches 120; Indels 62; Gaps 15;

Qy 12 LLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVP-P 70  
 :| || | | | ||| :||: ||||| ||||| :| ||| | | |  
 Db 5 VLSTFLLHLVLP--VALSLSTCSTLDMDQFMRKRIEAIHQILSKLKLTSPP--DYPEP 60

Qy 71 GPLPEAVLALYNSTRD---RVAGESVEPEPE-PEADYYAKEVTRVLM---VESGNQIYD 122  
 :| |::| |||| : : : | | : :| |||| : : | : | |  
 Db 61 DEVPPEVISIYNSTRDLLQEKASRRAAACERERSDEEYYAKEVYKIDMPSHLPSENAIPP 120

Qy 123 KFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----K 173  
 | | : | : : | :|| | : | | :| || :|||  
 Db 121 TFY-RPYFRIVRFVDVSTMEKNASN--LVKAEFRVFRQLQNPKEARVAEQRIELYQILKSKD 176

Qy 174 YSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC----- 225  
 : : ||: :: : ||||| |::|| : : ||:| || |  
 Db 177 LTSPTQRYIDSKVVKTRAEGEWLSFDVTDVQEWLHHKDRNLGFKISLHCPCTFVPSNN 236

Qy 226 ---DSKDNTLHVE---INGFNSGRRGDLATIHGMR-----PFLLLMATPLERAQHLHS 273  
 :| | :| : : || || : | ||| | | : |  
 Db 237 YIIPNKSEELEARFAGIDGTSTYASGDQKTIKSTRKKTSGKTPHLLLMLLPSYRLESQQS 296

Qy 274 SRHRRALDTNSYPYDVPDYASLALDTNYCFSSSTEKNCCVRQLYIDFRKDLGWKWIHEPKG 333  
 || : : ||| ||| : : |||:| ||||:| ||||| |||||  
 Db 297 SRRKKR-----ALDAAYCFRNVQDNCCLRPPLYIDFKRDLGWKWIHEPKG 340

Qy 334 YHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKV 393  
 | :||| | |||:| |||:| |||:| |||:| ||| | ||| | :||:| ||:  
 Db 341 YNANFCAGACPYLWSSDTQHTKVLSLYNTINPEASAPCCVSDLEPLTILYYIGNTPKI 400

Qy 394 EQLSNMIVRSCKCS 407  
 ||||| |||: |||||  
 Db 401 EQLSNMIVKSKCS 414

RESULT 8  
 O93449

ID O93449 PRELIMINARY; PRT; 382 AA.  
 AC O93449; Q91217;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta precursor.  
 GN TGF-BETA OR TGF.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEUKOCYTE;  
 RX MEDLINE=99242020; PubMed=10227481;  
 RA Daniels G.D., Secombes C.J.;  
 RT "Genomic organisation of rainbow trout, *Oncorhynchus mykiss* TGF-  
 RT BETA.";  
 RL Dev. Comp. Immunol. 23:139-147(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=LEUKOCYTE;  
 RX MEDLINE=98390168; PubMed=9722928;  
 RA Hardie L.J., Laing K.J., Daniels G.D., Grabowski P.S., Cunningham C.,  
 RA Secombes C.J.;  
 RT "Isolation of the first piscine transforming growth factor beta gene:  
 RT analysis reveals tissue specific expression and a potential regulatory  
 RT sequence in rainbow trout (*Oncorhynchus mykiss*).";  
 RL Cytokine 10:555-563(1998).  
 CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE  
 CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BLOOD LEUKOCYTES, KIDNEY  
 CC MACROPHAGES, BRAIN, GILL AND SPLEEN BUT NOT IN LIVER.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AJ007836; CAA07707.1; -.  
 DR EMBL; X99303; CAA67685.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 270  
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.  
 FT DISULFID 278 286 BY SIMILARITY.  
 FT DISULFID 285 348 BY SIMILARITY.  
 FT DISULFID 314 379 BY SIMILARITY.  
 FT DISULFID 318 381 BY SIMILARITY.  
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 237 237 N -> D (IN REF. 2).  
 FT CONFLICT 345 345 Q -> H (IN REF. 2).  
 FT CONFLICT 371 372 LS -> VP (IN REF. 2).

Query Match 38.1%; Score 826; DB 13; Length 382;  
Best Local Similarity 46.8%; Pred. No. 2.8e-65;  
Matches 184; Conservative 57; Mismatches 104; Indels 48; Gaps 14;

Qy	30	LSTCKTIDMELVKKRIEAIRGQILSKLRLASPP--SQGDVPPGGLPEAVLALYNSTRD	86
		:   :: :                      :   :  ::::     :	
Db	23	MSTCKSLDLELVKKRIEAIRGQILSKLRLPKEPEIDQEGDTE--EVPASLMSIYNSTVE	80
Qy	87	RVAGESVE-----PEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSEL	140
		::    :       :     :   :    :   :       :	
Db	81	-LSEEQVHTYIPSTQDAEEEA-YFAKEVHKFNMQSENT-----SKHQI--LFNMSEM	129
Qy	141	REAVPEPVLLSRAELRLL---RLKLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWL	196
		:    :          : :  :     : : :   :	
Db	130	RSVLGTDRLLSQAELRLLIKNHGLLDDSEQRLELYRGV-GDKARYLKSHFVSKEWANRWV	188
Qy	197	SFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFNSGRRGDLATI--HGMN	254
		: :   :  :      :  :      :   : :	
Db	189	SFDVTQTLNEWLQGAGEEQGFQLKLPCDCGKPMEEFRFKISGMNK-LRGNTETLAMKMPS	247
Qy	255	RPFLLLMTPLERAQHLHSSRHRRALDNTNSYPYDVPDYASLALDNTNYCFSSTEKNCCVRQ	314
		:  :   :  :         :        :   :	
Db	248	KPHILLMSLPVERHSQL-SSRKKRQTTE-----EIC-SDKSESCCVRK	289
Qy	315	LYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCV	374
		: :     : : : :	
Db	290	LYIDFRKDLGWKWIHEPTGYFANYCIGPCTYIWNTENKYSQVLALYKHHNPGASAPCCV	349
Qy	375	PQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS	407
		:     :          : :	
Db	350	PQVLEPLPIIYYVGRQHKVEQLSNMIVKSCRCS	382

Q9PTQ2

```

ID      Q9PTQ2                PRELIMINARY;          PRT;    376 AA.
AC      Q9PTQ2;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta precursor.
OS      Cyprinus carpio (Common carp).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Cyprinus.
OX      NCBI_TaxID=7962;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Yin Z., Kuang J.;
RT      "Molecular cloning of carp transforming growth factor beta 1.";
RL      Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC      -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE
CC      RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.
CC      -!- SUBUNIT: HOMODIMER. DISULFIDE-LINKED (BY SIMILARITY).

```



Db 325 VLALYKHHNPGASAOPCRVPQVLNPLPIFYVVGROHKVEQLSNMIVKTCKC 375

RESULT 10

Q9ERB7

ID Q9ERB7 PRELIMINARY; PRT; 399 AA.  
 AC Q9ERB7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Transforming growth factor-beta 2 (Fragment).  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ramesh G., Kondaiah P., Seshagiri P.B.;  
 RT "Differential expression and selective localization of transforming  
 RT growth factor-beta isoforms in the hamster uterus during estrous  
 RT cycle.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AY007214; AAG02247.1; -.  
 DR HSSP; P08112; 2TGI.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 399 399  
 SQ SEQUENCE 399 AA; 46078 MW; A6FF8E65EAFD5148 CRC64;

Query Match 36.0%; Score 780; DB 11; Length 399;  
 Best Local Similarity 42.6%; Pred. No. 3.8e-61;  
 Matches 179; Conservative 58; Mismatches 121; Indels 62; Gaps 15;

Qy 18 LLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVP-PGPLPEA 76  
 ||| | | ||||| :||| : ||||| ||||| ||||| : ||| | | :|  
 Db 4 LLHLVP--VALSLSTCSTLDMQFMRKRIEAIRGQILSKLKLTSPP--DYPEPDEVPP 59  
 Qy 77 VLALYNSTRD----RVAGESVEPEPE-PEADYYAKEVTRVLM--VESGNQIYDKFKGTP 128  
 |::| ||||| : : : | | : : ||||| : : | | | | | |  
 Db 60 VISIYNSTRDLLQEKASRRAAACERERSDEEYAKEVYKIDMPSHFPSENAIPPTFY-RP 118  
 Qy 129 HSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----KYSNDSW 179  
 : : | : : | : || : || | : | || : |||| : : :  
 Db 119 YFRIVRFDVSMMEKNASN---LVKAEFRVRLQNPKEARVAEQRIELYQILKSKDLTSPTQ 175  
 Qy 180 RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSC-----DSK 228  
 || : :: : ||||| | : || : : || : || | : |  
 Db 176 RYIDSKVVKTRAEGEWLSFDVTDVHEWLHHKDRNLGFKISLHCPCTFVPFNNNIIPNK 235

```

QY      229 DNTLHVE---INGFNSGRRGDLATIHGMNR-----PFLLLMATPLERAQHLHSSRRHRA 279
      |      | : | :      | ||      :      | |||| | | : | : | :
Db      236 SEELEARFAGIDGTSQHSSGHQETIKSTRKKNSGKTPHLLMLLPSYRLESQQSNRRKKR 295

QY      280 LDTNSYPYDVPDYASIALDNTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFC 339
      ||| ||| : : |||:| |||||:| ||||| |||||:| |||
Db      296 -----ALDAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKGYNANFC 339

QY      340 LGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNM 399
      | |||:| |||:| |||:| || |||:| ||| ||| | : ||| | ||| |||
Db      340 AGACPYLWSSDTQHTKVLSLYNTINPEASASPCCVSHDLEPLTILYYIGNTPKIEQLSNM 399

```

RESULT 11

Q99K17

```

ID   Q99K17      PRELIMINARY;          PRT;   362 AA.
AC   Q99K17;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to transforming growth factor, beta 3 (Fragment).
GN   TGFB3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; BC005513; AAH05513.1; -.
DR   HSSP; P10600; 1TGJ.
DR   MGD; MGI:98727; Tgfb3.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFB.
DR   InterPro; IPR001111; TGFB_N.
DR   InterPro; IPR003911; TGF_TGFB.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFB_propeptide; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFBETA.
DR   ProDom; PD000357; TGFB; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
FT   NON_TER      1      1
SQ   SEQUENCE      362 AA;  41486 MW;  0808E46180FDAE70 CRC64;

```

```

Query Match          34.0%;   Score 736.5;   DB 11;   Length 362;
Best Local Similarity 42.9%;   Pred. No. 2.6e-57;
Matches 164;   Conservative 51;   Mismatches 114;   Indels 53;   Gaps 12;

```

```

QY      58 RLASPPSQGDVPPGPLPEAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVL 112
      | : |||      :      : | ||||| |||      : : || |      : | : ||||| : :
Db      2  RVGSPPEPSVMT--HVPYQVLALYNSTRELL EEMHGEREEGCTQETSESEYYAKEIHKFD 59

```

Qy 113 MVE---SGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVE 165  
 |:: |:: || :: || |:: | || |:: |:: : |  
 Db 60 MIQGLAEHNELAVCPKGITSKVFR-FNVSSVEK---NGTNLFRAEFRVLVNPSSSKRTE 115

Qy 166 QHVELYQKYSND----SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSA 221  
 |::|::| |:: : | : ||||| |::| |:: | : |  
 Db 116 QRIELFQILRPDEHIAKQRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISI 175

Qy 222 HCSC-----DSKDNTLHVEINGFNS---GRRGDLATIHGM---NRPFLLLMATPL 265  
 || | :: : : | :: ||| : | |:: |  
 Db 176 HCPCHTFQPNGDILENVHEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLILMMIPP 235

Qy 266 ERAQHLHSSRRHRRALDTSYPYDVPDYASLALDTSNYCFSSTEKNCCVRQLYIDFRKDLGW 325  
 | : ||||| : |::| |::| |::|  
 Db 236 HRLDSPGQGSQRK-----KRALDTSNYCFRNLEENCCVRPLYIDFRQDLGW 280

Qy 326 KWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVY 385  
 ||:|||||:|||| |::| |::| |::| |::| |::| |::| |::|  
 Db 281 KKWHEPKGYANFCSGPCPYLRSADTTHTSTVLGLYNTLNPEASASPCCVPQDLEPLTILY 340

Qy 386 YVGRKPKVEQLSNMIVRSCKCS 407  
 |||| |::|  
 Db 341 YVGRTPKVEQLSNMVVKSCCKCS 362

# RESULT 12

Q98854

ID Q98854 PRELIMINARY; PRT; 361 AA.  
 AC Q98854;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta 2 (TGF-beta 2) (Fragment).  
 GN TGFB2.  
 OS Cyprinus carpio (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 OX NCBI\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEART;  
 RX MEDLINE=97354301; PubMed=9210595;  
 RA Sumathy K., Desai K.V., Kondaiah P.;  
 RT "Isolation of transforming growth factor-beta2 cDNA from a fish,  
 RT Cyprinus carpio by RT-PCR."  
 RL Gene 191:103-107(1997).  
 CC -!- FUNCTION: TGF-BETA 2 HAS SUPPRESSIVE EFFECTS ON INTERLEUKIN-2  
 CC DEPENDENT T-CELL GROWTH.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; U66874; AAB62983.1; -.  
 DR HSSP; P08112; 2TGI.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.



DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).  
 GN TGFB1.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LVG (SYR);  
 RX MEDLINE=93304479; PubMed=8317544;  
 RA Wong D.T., Donoff R.B., Yang J., Song B.Z., Matossian K., Nagura N.,  
 RA Elovic A., McBride J., Gallagher G., Todd R.;  
 RT "Sequential expression of transforming growth factors alpha and beta 1  
 RT by eosinophils during cutaneous wound healing in the hamster."  
 RL Am. J. Pathol. 143:130-142(1993).  
 RN [2]  
 RP SEQUENCE OF 26-115 FROM N.A.  
 RC STRAIN=SYRIAN; TISSUE=SPLEEN;  
 RX MEDLINE=98234044; PubMed=9573100;  
 RA Melby P.C., Tryon V.V., Chandrasekar B., Freeman G.L.;  
 RT "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and  
 RT analysis of cytokine mRNA expression in experimental visceral  
 RT leishmaniasis."  
 RL Infect. Immun. 66:2135-2142(1998).  
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF  
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1  
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND  
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.  
 DR EMBL; X60296; CAA42838.1; -.  
 DR EMBL; AF046214; AAC40099.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT PROPEP <1 18  
 FT CHAIN 19 130 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 25 34 BY SIMILARITY.  
 FT DISULFID 33 96 BY SIMILARITY.  
 FT DISULFID 66 129 BY SIMILARITY.  
 FT DISULFID 95 95 INTERCHAIN (BY SIMILARITY).  
 FT CONFLICT 93 93 G -> S (IN REF. 2).  
 SQ SEQUENCE 130 AA; 14997 MW; 8B41DD6CF39CCA77 CRC64;

Query Match 31.9%; Score 691.5; DB 11; Length 130;  
 Best Local Similarity 86.4%; Pred. No. 6.3e-54;  
 Matches 127; Conservative 0; Mismatches 3; Indels 17; Gaps 1;

Qy	261	MATPLERAQHLHSSRRHRRALDNTSYDPYDVPDYASLALDNTNYCFSSTEKNCCVRQLYIDFR	320
Db	1	MATPLERAQHLQSSRHRR-----ALDNTNYCFSSTEKNCCVRQLYIDFR	43
Qy	321	KDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEP	380
Db	44	KDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAGPCCVPQALEP	103
Qy	381	LPIVYYVGRKPKVEQLSNMIVRSCKCS	407
Db	104	LPIVYYVGRKPKVEQLSNMIVRSYKCS	130

RESULT 14

095N80

```

ID      Q95N80          PRELIMINARY;          PRT;    124 AA.
AC      Q95N80;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta 1 (Fragment).
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Fonfara S., Groene A., Baumgaertner W.;
RT      "Sequence of canine transforming growth factor beta 1 mRNA in DH82-
RT      cells.";
RL      Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AF349538; AAK54072.1; -.
DR      InterPro; IPR001839; TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
FT      NON_TER          1          1
FT      NON_TER          124        124
SQ      SEQUENCE    124 AA;  14329 MW;  21D185218E5556DB CRC64;

```

Query Match 31.1%; Score 674.5; DB 6; Length 124;  
Best Local Similarity 87.2%; Pred. No. 1.9e-52;  
Matches 123; Conservative 0; Mismatches 1; Indels 17; Gaps 1;

Qy	264	PLERAQHLHSSRRRALD	TNSYPYDVPDYAS	LALD	TNYCFSSTEKNCCVRQ	LYIDFRKDL	323
Db	1	PLERAQHLHSSRQRR	-----	ALD	TNYCFSSTEKNCCVRQ	LYIDFRKDL	43
Qy	324	GWKWIHEPKGYHANF	CLGPCPYIWSLDTQ	YSKVLALYNQHNP	GASAAPCCVPQ	ALEPLPI	383
Db	44	GWKWIHEPKGYHANF	CLGPCPYIWSLDTQ	YSKVLALYNQHNP	GASAAPCCVPQ	ALEPLPI	103
Qy	384	VYYVGRKPKVEQL	SNMIVRSC	404			
Db	104	VYYVGRKPKVEQL	SNMIVRSC	124			

RESULT 15

O02730

ID O02730 PRELIMINARY; PRT; 112 AA.  
AC O02730; O97501;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).  
GN TGFBI OR TGF-BETA-1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Taylor T.K., James E.R., McGonigle S., Yoho E.R.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE OF 2-99 FROM N.A.  
RA Inoue K., Kawabe Y., Kodama T.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS  
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF  
CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1  
CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND  
CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR EMBL; AF000133; AAB53806.1; -.  
DR EMBL; AB020217; BAA36950.1; -.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein.  
FT NON\_TER 1 1  
FT CHAIN 1 112 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 7 16 BY SIMILARITY.  
FT DISULFID 15 78 BY SIMILARITY.  
FT DISULFID 44 109 BY SIMILARITY.  
FT DISULFID 48 111 BY SIMILARITY.  
FT DISULFID 77 77 INTERCHAIN (BY SIMILARITY).  
FT CONFLICT 2 3 LD -> FS (IN REF. 2).  
FT CONFLICT 85 92 PLPIVYYV -> ATAHRTTTL (IN REF. 2).  
SQ SEQUENCE 112 AA; 12795 MW; 53C5B7D46355A6F3 CRC64;

Query Match 29.4%; Score 638; DB 6; Length 112;  
Best Local Similarity 100.0%; Pred. No. 3.1e-49;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
Qy      296 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 60

Qy      356 VLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 407
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112
```

Search completed: October 28, 2003, 09:12:27  
Job time : 33.6306 secs